

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 22, 2004, 12:46:56 ; Search time 48 seconds  
(without alignments)  
1572.453 Million cell updates/sec

Title: US-09-978-191a-506  
Perfect score: 1505  
Sequence: 1 MRGSEVLLMWLLVAVGGT.....SEQISFLEQLGSCCKKDS 273

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1133595 seqs, 276475211 residues

Total number of hits satisfying chosen parameters: 1133595

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA:  
1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep:  
2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep:  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep:  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep:  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep:  
6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep:  
7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep:  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep:  
9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep:  
10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep:  
11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep:  
12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep:  
13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep:  
14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep:  
15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep:  
16: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep:  
17: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep:  
18: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1505	100.0	273	9	US-09-790-264-10
2	1505	100.0	273	9	US-09-978-295A-506
3	1505	100.0	273	9	US-09-978-295A-508
4	1505	100.0	273	9	US-09-978-697-506
5	1505	100.0	273	9	US-09-978-697-508
6	1505	100.0	273	9	US-09-978-192A-506
7	1505	100.0	273	9	US-09-978-192A-508
8	1505	100.0	273	9	US-09-999-832A-506
9	1505	100.0	273	9	US-09-999-832A-508
10	1505	100.0	273	10	US-09-978-189-506
11	1505	100.0	273	10	US-09-978-189-508
12	1505	100.0	273	10	US-09-978-608A-506
13	1505	100.0	273	10	US-09-978-608A-508
14	1505	100.0	273	10	US-09-978-585A-506
15	1505	100.0	273	10	US-09-978-585A-508

ALIGNMENTS

RESULT 1

US-09-790-264-10

; Sequence 10, Application US/09790264

; GENERAL INFORMATION:

; APPLICANT: Holtzman, Douglas A.

; APPLICANT: Goodearl, Andrew D.J.

; APPLICANT: McCarthy, Sean A.

; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING

; TITLE OF INVENTION: PROGNASTIC, DIAGNOSTIC, PREVENTIVE,

; TITLE OF INVENTION: THERAPEUTIC, AND OTHER

; FILE REFERENCE: 07334-322001

; CURRENT APPLICATION NUMBER: US/09/790,264

; CURRENT FILING DATE: 2001-02-21

; PRIOR APPLICATION NUMBER: US 09/065,661

; PRIOR FILING DATE: 1998-04-23

; PRIOR APPLICATION NUMBER: US 09/298,531

; PRIOR FILING DATE: 1999-04-23

; PRIOR APPLICATION NUMBER: US 09/065,363

; PRIOR FILING DATE: 1998-04-23

; PRIOR APPLICATION NUMBER: US 09/337,930

; PRIOR FILING DATE: 1999-06-22

; PRIOR APPLICATION NUMBER: US 09/102,705

; PRIOR FILING DATE: 1998-06-22

; PRIOR APPLICATION NUMBER: US 09/363,630

; PRIOR FILING DATE: 1999-07-29

; PRIOR APPLICATION NUMBER: US 09/124,538

; PRIOR FILING DATE: 1998-07-29

; NUMBER OF SEQ ID NOS: 68

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 10

; LENGTH: 273

; TYPE: PRT

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: SIGNAL

; LOCATION: (1)...(22)

; US-09-790-264-10

Seqs & Alignments

Query Match 100.0%; Score 1505; DB 9; Length 273;  
Best Local Similarity 100.0%; Pred. No. 1.1e-113;  
Matches 273; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRGSOEVLMLLLVLAUGTGHAYRPGRRVCVAHAGDPVSEFVQRYOPFLTTCDGHR 60  
DB 1 MRGSOEVLMLLLVLAUGTGHAYRPGRRVCVAHAGDPVSEFVQRYOPFLTTCDGHR 60

QY 61 ACSTYRTIYAYRSPGLAPARPRYACCPGWKTSGLPGAGAAICQPPCRNGGSCVQP 120  
DB 61 ACSTYRTIYAYRSPGLAPARPRYACCPGWKTSGLPGAGAAICQPPCRNGGSCVQP 120

QY 121 GRCPAGWRGDTQSDVDECSARRGCPQRCINTAGSYWCQWEGHSLGADGTLCPVKG 180  
DB 121 GRCPAGWRGDTQSDVDECSARRGCPQRCINTAGSYWCQWEGHSLGADGTLCPVKG 180

QY 181 GPRVAPNPTGVDGAMKEEVORLQSRVDLLBEKQLVLAPLHSLASQALEHGLPDPGSL 240  
DB 181 GPRVAPNPTGVDGAMKEEVORLQSRVDLLBEKQLVLAPLHSLASQALEHGLPDPGSL 240

QY 241 VHSFOQLGRIDSLSEHSFLEEQGSCSKDS 273  
DB 241 VHSFOQLGRIDSLSEHSFLEEQGSCSKDS 273

## RESULT 2

US-09-978-295A-506  
Sequence 506, Application US/09978295A  
Patent No. US2002015606A1  
GENERAL INFORMATION:  
APPLICANT: Ashkenazi, Avi  
APPLICANT: Baker Kevin P.  
APPLICANT: Botstein, David  
APPLICANT: Desnoyers, Luc  
APPLICANT: Eaton, Dan  
APPLICANT: Ferrara, Napoleon  
APPLICANT: Filvaroff, Ellen  
APPLICANT: Fong, Sherman  
APPLICANT: Gao, Wei-Qiang  
APPLICANT: Gerber, Hanspeter  
APPLICANT: Gerritsen, Mary E.  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Grimaldi, J. Christopher  
APPLICANT: Gurney, Austin L.  
APPLICANT: Hillan, Kenneth J.  
APPLICANT: Kljavin, Ivar J.  
APPLICANT: Kuo, Sophia S.  
APPLICANT: Napier, Mary A.  
APPLICANT: Pan, James  
APPLICANT: Paoli, Nicholas F.  
APPLICANT: Roy, Margaret Ann  
APPLICANT: Shelton, David L.  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Williams, P. Mickey  
APPLICANT: Wood, William I.  
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
FILE OF INVENTION: Acids Encoding the Same  
FILE REFERENCE: P2630P1C11  
CURRENT APPLICATION NUMBER: US/09/978,295A  
CURRENT FILING DATE: 2001-10-15  
PRIOR APPLICATION NUMBER: 09/918585  
PRIOR FILING DATE: 2001-07-30  
PRIOR APPLICATION NUMBER: 60/062250  
PRIOR FILING DATE: 1997-10-17  
PRIOR APPLICATION NUMBER: 60/064249  
PRIOR FILING DATE: 1997-11-03  
PRIOR APPLICATION NUMBER: 60/065311  
PRIOR FILING DATE: 1997-11-13  
PRIOR APPLICATION NUMBER: 60/066364  
PRIOR FILING DATE: 1997-11-21

APPLICANT: Ashkenazi, Avi  
APPLICANT: Baker Kevin P.  
APPLICANT: Botstein, David  
APPLICANT: Desnoyers, Luc  
APPLICANT: Eaton, Dan  
APPLICANT: Ferrara, Napoleon  
APPLICANT: Filvaroff, Ellen  
APPLICANT: Fong, Sherman  
APPLICANT: Gao, Wei-Qiang  
APPLICANT: Gerber, Hanspeter  
APPLICANT: Gerritsen, Mary E.  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Grimaldi, J. Christopher  
APPLICANT: Gurney, Austin L.  
APPLICANT: Hillan, Kenneth J.  
APPLICANT: Kljavin, Ivar J.  
APPLICANT: Kuo, Sophia S.  
APPLICANT: Napier, Mary A.  
APPLICANT: Pan, James  
APPLICANT: Paoli, Nicholas F.  
APPLICANT: Roy, Margaret Ann  
APPLICANT: Shelton, David L.  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Williams, P. Mickey  
APPLICANT: Wood, William I.  
PRIOR APPLICATION NUMBER: 60/077450  
PRIOR FILING DATE: 1998-03-10  
PRIOR APPLICATION NUMBER: 60/077632  
PRIOR FILING DATE: 1998-03-11  
PRIOR APPLICATION NUMBER: 60/077641  
PRIOR FILING DATE: 1998-03-11  
PRIOR APPLICATION NUMBER: 60/077649  
PRIOR FILING DATE: 1998-03-11  
PRIOR APPLICATION NUMBER: 60/077791  
PRIOR FILING DATE: 1998-03-12  
PRIOR APPLICATION NUMBER: 60/078004  
PRIOR FILING DATE: 1998-03-13  
PRIOR APPLICATION NUMBER: 60/078886  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/078936  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/078910  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/078939  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/079294  
PRIOR FILING DATE: 1998-03-25  
PRIOR APPLICATION NUMBER: 60/079656  
PRIOR FILING DATE: 1998-03-26  
PRIOR APPLICATION NUMBER: 60/079664  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079689  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079663  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079728  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079786  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079920  
PRIOR FILING DATE: 1998-03-30  
PRIOR APPLICATION NUMBER: 60/079923  
PRIOR FILING DATE: 1998-03-30  
PRIOR APPLICATION NUMBER: 60/080105  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/080107  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/080165  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/080194  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/080327  
PRIOR FILING DATE: 1998-04-01  
PRIOR APPLICATION NUMBER: 60/080328  
PRIOR FILING DATE: 1998-04-01  
PRIOR APPLICATION NUMBER: 60/080333  
PRIOR FILING DATE: 1998-04-01  
PRIOR APPLICATION NUMBER: 60/080334  
PRIOR FILING DATE: 1998-04-01  
PRIOR APPLICATION NUMBER: 60/081070  
PRIOR FILING DATE: 1998-04-08  
PRIOR APPLICATION NUMBER: 60/081049  
PRIOR FILING DATE: 1998-04-08  
PRIOR APPLICATION NUMBER: 60/081071  
PRIOR FILING DATE: 1998-04-08  
PRIOR APPLICATION NUMBER: 60/081195  
PRIOR FILING DATE: 1998-04-08  
PRIOR APPLICATION NUMBER: 60/081203  
PRIOR FILING DATE: 1998-04-09  
PRIOR APPLICATION NUMBER: 60/081229  
PRIOR FILING DATE: 1998-04-09  
PRIOR APPLICATION NUMBER: 60/081955  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081817  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081819  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081952

F;702-704/Region: cell attachment (R-G-D) motif

F;714-750/Domain: EGF homology <EG3>  
F;762-800/Domain: EGF homology <EG4>  
F;806-839/Domain: EGF homology <EG5>  
F;849-919/Domain: thyroglobulin type I repeat homology <THV1>  
F;990-1032/Domain: LDL receptor WYTD-containing repeat homology <YML>  
F;1033-1075/Domain: LDL receptor WYTD-containing repeat homology <YML>  
F;1076-1120/Domain: LDL receptor WYTD-containing repeat homology <YML>  
F;1121-1160/Domain: LDL receptor WYTD-containing repeat homology <YML>  
F;1161-1197/Domain: LDL receptor WYTD-containing repeat homology <YML>  
F;1212-1243/Domain: EGF homology <EG6>  
F;289, 296/Binding site: sulfate (tyr) (covalent) #status predicted  
F;729, 819/Modified site: erythro-beta-hydroxyasparagine (Asn) #status predicted  
F;756/Cross-link: isopeptide (Gln) (interchain to Lys N6-amino of laminin) #status predicted  
F;1137/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 11.5%; Score 173.5; DB 1; Length 1247;  
Best Local Similarity 25.8%; Pred. No. 4.6e-05;  
Matches 80; Conservative 23; Mismatches 98; Indels 109; Gaps 18;

QY 20 TEHAYRPGRRV-----CAVRHG-----DPVSESFVQRYVQPLTTCDGHRACSTYR 66  
Db 681 TNAACRPGRTPTCECSIGFRDGTCTVDICSE-----QP-SVCGSHITCNH- 730  
QY 67 TIYRTAYRSPGLAPRPIYACPGWKRRTSGLFGACGAICOPP-----CRNG----- 114  
Db 731 -----PGTFRCECVGQFSD--EGTCVAVDQRFNYCETGLHNCIDIPQR 774  
QY 115 GSCVQPG-----RCRCAGHRGD--TCOSDVDCSARRGCPORCINTAGSYWCOCWEGHS 168  
Db 775 AQCIYTGSSSYTCLPLGFGSGDQACQ-DVDEBCQSRCHPDFAFCYNTPGSTCCQKPGY- 832  
QY 169 LSADGTLCPV-----KGGPRVAPNPTG----- 191  
Db 833 -QSGDFRCVPGVEKTECQHERHILGAAGATDPQRIPLGLFVPCDAGHYAFTQCHG 891  
QY 192 -----VDSAMKEEVQRLQSRVDLLEKQLVLAPLH-SLASQALEHGLPPGSLVHS 243  
Db 892 STGYCNCVDRDGR-EVEGTRTRFGMTFPCPLSTVAPPIHQGPVFTAVIPLP-PGTHLL-- 947  
QY 244 FQQLGRIDSL 253  
Db 948 FAQTGKIERL 957

## RESULT 15

A24420  
notch protein - fruit fly (Drosophila melanogaster)  
N;Alternate names: neurogenic repetitive locus protein  
C;Species: Drosophila melanogaster  
C;Date: 10-Sep-1999 #sequence revision 10-Sep-1999 #text\_change 10-Sep-1999  
C;Accession: A24420; A24768; S093358; A05267  
R;Kidd, S.; Kelley, M.R.; Young, M.W.  
Mol. Cell. Biol. 6, 3094-3108, 1986  
A;Reference number: A24420; MUID: 97064624; PMID: 3097517  
A;Accession: A24420  
A;Molecule type: DNA  
A;Residues: 1-2703 <KID>  
A;Cross-references: GB:K03508; NID: g157991; PID: AAA28725.1; PID: g157993  
R;Wharton, K.A.; Johansen, K.M.; Xu, T.; Artavanis-Tsakonas, S.  
Cell 43, 567-581, 1985  
A;Reference number: A24768; MUID: 86079539; PMID: 3935325  
A;Accession: A24768  
A;Molecule type: mRNA  
A;Residues: 1-48, 'I', 50-118, 'R', 120-230, 'I', 232-256, 'N', 258-266, 'A', 268-872, 'R', 874-958,  
R;Tautz, D.  
Nucleic Acids Res. 17, 6463-6471, 1989  
A;Title: Hypervariability of simple sequences as a general source for polymorphic DNA ma  
A;Reference number: S093358; MUID: 89385974; PMID: 2780284  
A;Accession: S093358  
A;Molecule type: DNA  
A;Residues: 2505-2551, 'QQQ', 2552-2576, 'E', 2578-2604 <TAU>

R;Wharton, K.A.; Yedvobnick, B.; Finnerty, V.G.; Artavanis-Tsakonas, S.  
Cell 40, 55-62, 1985  
A;Title: opa: a novel family of transcribed repeats shared by the Notch locus and other.  
A;Reference number: A05267; MUID: 85099329; PMID: 2981631

A;Accession: A05267

A;Molecule type: DNA

A;Residues: 2504-2576, 'E', 2578-2611 <WHA2>

C;Genetics:

A;Gene: notch; opa

A;Cross-references: FlyBase: FBgn0004647

A;Map position: 8.96-9.36

A;Introns: 53/3; 84/3; 171/3; 240/3; 283/3; 2333/3; 2436/3; 2588/3

C;Superfamily: notch protein; ankyrin repeat homology; EGF homology

C;Keywords: differentiation; tandem repeat; transmembrane protein

F;27-43/Domain: transmembrane #status predicted <TM1>

F;297-328/Domain: EGF homology <EGX1>

F;530-561/Domain: EGF homology <EGF1>

F;568-599/Domain: EGF homology <EGF2>

F;1064-1095/Domain: EGF homology <EGX2>

F;1187-1218/Domain: EGF homology <EGX3>

F;1746-1762/Domain: transmembrane #status predicted <TM2>

F;1950-1982/Domain: ankyrin repeat homology <AN1>

F;1983-2015/Domain: ankyrin repeat homology <AN2>

F;1998-2004/Domain: transmembrane #status predicted <TM3>

F;2017-2049/Domain: ankyrin repeat homology <AN3>

F;2050-2082/Domain: ankyrin repeat homology <AN4>

F;2083-2115/Domain: ankyrin repeat homology <AN5>

F;2538-2568/Region: glutamine-rich

F;2538-2568/Domain: neurogenic repetitive element #status predicted <OPA>

Query Match 11.5%; Score 173; DB 1; Length 2703;

Best Local Similarity 26.7%; Pred. No. 9.9e-05;  
Matches 59; Conservative 19; Mismatches 67; Indels 76; Gaps 13;

QY 18 GGTEHAYRPGRRVCAVRAH--GDPVSESFVQRYVQPLTTCDGHRACSTYRTIYRTAYRR 75  
Db 70 GGTCTVQLNGKTYCACDSHYVGD-----YCEHRNFCNSNR-CQNGTCCQVTFRNG 118  
QY 76 SPGLAPRPIYAC-CP-GWKRT---SGLPGACGAIC----- 107  
Db 119 HPGI-----SCKPLGFDSELCIEIAVPNACDHVTCLNGTCLQKLTLEYTCACANGYT 171  
QY 108 -----QPPCRNGGSCV-----QPGRCRCPAGWRGDTCCSDVDECSA---RRGGC 148  
Db 172 GERCETKNLCASSPCRNATCTALAGSSSFTSCPPGFTGDTCTSYDIEECQSNPKYGG- 230  
QY 149 PQRCTAGSYWCOCWEGHLSADGTLCPVKGPPRVPAPNP 189  
Db 231 --TCVNTHGSYQCMCPGTGT----GKDCDTKYK- -GSPSP 263

Search completed: April 22, 2004, 12:47:22  
Job time : 22 secs

## S78549

notch3 protein - human  
C:Species: Homo sapiens (man)  
C:Date: 24-Jul-1998 #sequence\_revision 24-Jul-1998 #text\_change 08-Sep-2002  
C:Accession: S78549; S71825  
R:Joutel, A.; Tournier-Lasserre, E.  
submitted to the EMBL Data Library, April 1997  
A:Reference number: S78549  
A:Accession: S78549  
A:Molecule type: mRNA  
A:Residues: 1-2321 <JOU1>  
A:Cross-references: EMBL:U97669; NID:92668591; PIDN:AB91371.1; PID:92668592  
R:Joutel, A.; Corpechot, C.; Ducros, A.; Vahedi, K.; Chabriat, H.; Mouton, P.; Alamowicz, M.M.; Weissbach, J.; Bach, J.F.; Bousser, M.G.; Tournier-Lasserre, E.  
Nature 383, 707-710, 1996  
A:Title: Notch3 mutations in CADASIL, a hereditary adult-onset condition causing stroke  
A:Reference number: S71825; MUID:97032728; PMID:8878478  
A:Accession: S71825  
A:Status: nucleic acid sequence not shown  
A:Molecule type: DNA  
A:Residues: 67-113;138-194;268-333; 'G', 335-346;536-613;716-765;1240-1279;1815-1868 <JOU2>  
A:Cross-references: EMBL:U97669  
C:Genetics:  
A:Map position: 19p13.1

C:Function:  
A:Description: may be involved in pathogenesis of CADASIL, causing a type of stroke and C:Superfamily: notch protein; ankyrin repeat homology; EGF homology  
C:Keywords: tandem repeat; transmembrane protein  
F:123-155/Domain: EGF homology <EGX1>  
F:162-194/Domain: EGF homology <EGF1>  
F:240-271/Domain: EGF homology <EGX2>  
F:318-343/Domain: EGF homology <EGF>  
F:473-504/Domain: EGF homology <EGX3>  
F:853-884/Domain: EGF homology <EGF3>  
F:928-959/Domain: EGF homology <EGX4>  
F:1838-1870/Domain: ankyrin repeat homology <AN1>  
F:1871-1903/Domain: ankyrin repeat homology <AN2>  
F:1905-1937/Domain: ankyrin repeat homology <AN3>  
F:1938-1970/Domain: ankyrin repeat homology <AN4>  
F:1971-2003/Domain: ankyrin repeat homology <AN5>

## Query Match 11.7%; Score 175.5; DB 2; Length 2321;

Best Local Similarity 34.5%; Pred. No. 5.7e-05;  
Matches 50; Conservative 14; Mismatches 44; Indels 37; Gaps 11;

QY 56 CDHPRACSTYRTIYTRSPGLAPAPRYAC-CPGWKR--TSGLPACGAAICQPPCR 112  
DB 87 CAGRGVCOS-----SVVAGTARFSCRCRFRGPDCLPDC----LSSPCA 129  
QY 113 NGGSC-VOP-GR--CRCPAGWRGDTCSQSDVDEC-----SARRGGCPQRCINTAGSYWCQW 164  
DB 130 HGARCSVGDGRFLCSCPFGVQGRSCSDVDECVRGECRHGG---TCLNTPGSRFCQCP 186  
QY 165 EGHLSADGTLCPVKGGRPRVADNP 189  
DB 187 AGYT----GPLCENPAPV--CAPSP 205

## RESULT 13

nel protein - chicken  
C:Species: Gallus gallus (chicken)  
C:Date: 24-Feb-1995 #sequence\_revision 24-Feb-1995 #text\_change 13-Aug-1999  
C:Accession: A38963; JP0076  
R:Matuhashi, S.; Noji, S.; Koyama, E.; Myokai, F.; Ohuchi, H.; Taniguchi, S.; Hori, K.  
Dev. Dyn. 203, 212-222, 1995  
A:Title: New gene, nel, encoding a Mr 93K protein with EGF-like repeats is strongly expressed in the developing chick  
A:Reference number: A38963; MUID:95383734; PMID:7655083  
A:Accession: A38963  
A:Molecule type: mRNA  
A:Residues: 1-835 <MAT>  
A:Cross-references: DBBJ:D45365

A:Experimental source: 9-day embryo  
R:Matuhashi, S.; Noji, S.; Koyama, E.; Myokai, F.; Ohuchi, H.; Taniguchi, S.; Hori, K.  
submitted to JIPID, January 1995  
A:Description: A new gene, nel, encoding a Mr 93K protein with EGF-like repeats is strongly expressed in the developing chick  
A:Reference number: JP0076  
A:Accession: JP0076  
A:Molecule type: mRNA  
A:Residues: 1-835 <MA2>  
A:Cross-references: DBBJ:D45365  
A:Experimental source: 9-day embryo  
C:Superfamily: von Willebrand factor type C repeat homology; EGF homology  
F:273-333/Domain: von Willebrand factor type C repeat homology <VWC>  
F:395-592/Region: EGF-like repeats  
F:444-480/Domain: EGF homology <EGF1>  
F:486-521/Domain: EGF homology <EGF>  
F:525-552/Domain: EGF homology <EGF2>

## Query Match 11.6%; Score 174; DB 2; Length 835;

Best Local Similarity 42.9%; Pred. No. 3e-05;  
Matches 30; Conservative 11; Mismatches 27; Indels 2; Gaps 1;

QY 100 GAGCAATCOPPRNGSCVOPGRCPAGWRGDTCSQSDVDECARRGGCPQR--CINTAG 157  
DB 518 GTVCAFAFKDGRNGGACIASNVCACPQPGFTGSPCETDIDECSDGFGVQCDSPANCINLP 577

QY 158 SYWCQCEGH 167

DB 578 WYHCECRDGY 597

## RESULT 14

MMHUND  
Nidogen precursor - human  
N:Alternate names: entactin  
C:Species: Homo sapiens (man)  
C:Date: 30-Sep-1991 #sequence\_revision 30-Sep-1991 #text\_change 20-Oct-2000  
C:Accession: A33322; A32437; A61367  
R:Nagayoshi, T.; Sanborn, D.; Hickok, N.J.; Olsen, D.R.; Fazio, M.J.; Chu, M.L.; Knowlton, D. 581-594, 1989  
A:Title: Human nidogen: complete amino acid sequence and structural domains deduced from cDNA cloning, cellular expression, and mapping of the gene to chromosome 8  
A:Reference number: A33322; MUID:90091745; PMID:2574658  
A:Accession: A33322  
A:Molecule type: mRNA  
A:Residues: 1-1247 <NAG>  
A:Cross-references: EMBL:M30269  
R:Olsen, D.R.; Nagayoshi, T.; Fazio, M.; Mattei, M.G.; Pasasege, E.; Weil, D.; Timpl, R.  
Am. J. Hum. Genet. 44, 876-885, 1989  
A:Title: Human nidogen: cDNA cloning, cellular expression, and mapping of the gene to chromosome 8  
A:Reference number: A32437; MUID:89270475; PMID:2471408  
A:Accession: A32437  
A:Molecule type: mRNA  
A:Residues: 667-1247 <OLS>  
A:Cross-references: EMBL:M27445; NID:9602466; PIDN:AA57261.1; PID:9602467  
A:Note: the authors translated the codon AAG for residue 966 as Cys  
R:Fazio, M.J.; O'Leary, J.; Kaeheari, V.M.; Chen, Y.Q.; Saitta, B.; Uitto, J.  
J. Invest. Dermatol. 97, 281-285, 1991  
A:Title: Human nidogen gene: structural and functional characterization of the 5'-flanking region  
A:Reference number: A61367; MUID:91302882; PMID:1906509  
A:Accession: A61367  
A:Molecule type: DNA  
A:Residues: 1-28 <FAZ>  
A:Comment: This protein is a basement membrane glycoprotein that forms a complex with nidogen  
C:Genetics:  
A:Gene: GDB:NID  
A:Cross-references: GDB:120236; OMIM:131390  
A:Map position: 1q43-1q43  
C:Superfamily: nidogen; EGF homology; LDL receptor YWTD-containing repeat homology; thyroglobulin; basement membrane; beta-hydroxyasparagine; calcium binding; cell binding; protein  
F:1-28/Domain: signal sequence #status predicted <SIG>  
F:29-1247/Product: nidogen #status predicted <NAT>  
F:390-425/Domain: EGF homology <EG1>  
F:672-708/Domain: EGF homology <EG2>

Db 504 CQPGVGN---GTICKAFCEGCRYGTCVAPNKCVCPSGFTGTHCEKIDCEACGFE 559  
QY 148 CP--ORCINTAGSYWCQWEGH-----SLSADGTLGV 177  
Db 560 CHNYSRCVNLPGWHCECRSGFHDGTYSLSGESCI 595

RESULT 10  
15476  
growth potentiating factor - rat  
C:Species: Rattus sp. (rat)  
C>Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 20-Apr-2001  
C:Accession: I55476  
R:Nakano, T.; Higashino, K.; Kikuchi, N.; Kishino, J.; Nomura, K.; Fujita, H.; Ohara, O.  
J. Biol. Chem. 270, 5702-5705, 1995  
A:Title: Vascular smooth muscle cell-derived, GliA-containing growth-potentiating factor  
A:Reference number: I55476; MUID:95197586; PMID:7890695  
A:Accession: I55476  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-674 <RES>  
A:Cross-references: GB:D42148; NID:G1526567; PIDN:BAA07719.1; PID:G993402  
C:Superfamily: plasma protein S; EGF homology; GliA domain homology; laminin G repeat homology  
F:29-99/Domain: GliA domain homology <GLA>  
F:117-150/Domain: EGF homology <EG1>  
F:157-192/Domain: EGF homology <EG2>  
F:198-233/Domain: EGF homology <EG3>  
F:239-274/Domain: EGF homology <EG4>  
F:308-667/Domain: sex hormone-binding globulin homology <SHB>  
F:318-470/Domain: laminin G repeat homology <LGR>

Query Match 11.8%; Score 178; DB 2; Length 674;  
Best Local Similarity 30.4%; Pred. No. 1.3e-05;  
Matches 48; Conservative 8; Mismatches 66; Indels 36; Gaps 6;  
QY 46 QRYVQFLTTCDGH-----RACSTYRTYTYARRSGLAPAPRYACC----- 89  
Db 44 RRAYVFEAKQGHLEECVCSKEA--REVPEDETDYFVPRYQECMRKYGREP 101  
QY 90 --PGWR--TSLPGAC-----GAAICOPFCRNGSCVQGRCPAGWGTGDCQSDV 138  
Db 102 KNPENFATVKNLDPDQTPNCPDKKGTQLCQDLGN-----FFCLCKDGMGRLCDKDV 154  
QY 139 DECSARRGGQPCRCINTAGSYWCQWEGHSLSDGTLG 176  
Db 155 NECSQKNGGSGVCHNKPQSFQACGHSFSLQSDNKS 192

RESULT 11  
A47221  
fibrillin 1 precursor - human (fragment)  
C:Species: Homo sapiens (man)  
C>Date: 02-Jun-1995 #sequence\_revision 25-Apr-1997 #text\_change 24-Nov-2003  
C:Accession: A47221; I54355; S17064; I59574; S17062; S62111; A34198  
R:Corson, G.M.; Chabberg, S.C.; Dietz, H.C.; Charbonneau, N.L.; Sakai, L.Y.  
Genomics 17, 476-484, 1993  
A:Title: Fibrillin binds calcium and is coded by cDNAs that reveal a multidomain structure  
A:Reference number: A47221; MUID:94010947; PMID:7691719  
A:Accession: A47221  
A:Molecule type: mRNA  
A:Residues: 1-337, 'T', 339-1029 <COR>  
A:Cross-references: GB:X63556  
R:Periera, L.V.; D'Alessio, M.; Ramirez, F.; Lynch, J.; Sykes, B.; Pangilinan, T.; Bonad  
Hum. Mol. Genet. 2, 961-968, 1993  
A:Title: Genomic organization of the sequence coding for fibrillin, the defective gene F  
A:Reference number: I54355; MUID:93372860; PMID:8364578  
A:Accession: I54355  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 132-3002 <PER>  
A:Cross-references: GB:U13923; NID:G306745; PIDN:AAB02036.1; PID:G306746  
R:Maslen, C.L.; Corson, G.M.; Maddox, B.K.; Glanville, R.W.; Sakai, L.Y.  
Nature 352, 334-337, 1991

A:Title: Partial sequence of a candidate gene for the Marfan syndrome.  
A:Reference number: S17064; MUID:91304568; PMID:1852207  
A:Accession: S17064  
A:Molecule type: mRNA  
A:Residues: 1030-3002 <MAS>  
A:Cross-references: EMBL:X63556  
R:Dietz, H.C.; Valle, D.; Francomano, C.A.; Kendzior, R.J.  
Science 259, 680-683, 1993  
A:Title: The skipping of constitutive exons in vivo induced by nonsense mutations.  
A:Reference number: I59574; MUID:93157831; PMID:8430317  
A:Accession: I59574  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 2217-2288, 'I', 2290-2325 <RES>  
A:Cross-references: GB:S54426; NID:G264860; PIDN:AAB2544.1; PID:G264861  
R:Lee, B.; Godfrey, M.; Vitale, E.; Horii, M.G.; Safarazi, M.; Tsipouras,  
Nature 352, 330-334, 1991  
A:Title: Linkage of Marfan syndrome and a phenotypically related disorder to two diffe  
A:Reference number: S17062; MUID:91304567; PMID:1852206  
A:Accession: S17062  
A:Molecule type: mRNA  
A:Residues: 'VLVTVFIFLSYNKML', 944-1444 <LEE1>  
A:Cross-references: EMBL:X62008; NID:G31398; PIDN:CAB56534.1; PID:G5924015  
A:Accession: S62111  
A:Molecule type: protein  
A:Residues: 1166-1176, 'X', 1178-1180, 'D', 1182-1185 <LEE2>  
R:Maddox, B.K.; Sakai, L.Y.; Keene, D.R.; Glanville, R.W.  
J. Biol. Chem. 264, 21381-21385, 1989  
A:Title: Connective tissue microfibrils. Isolation and characterization of three large  
A:Reference number: A34198; MUID:90078246; PMID:2512293  
A:Accession: A34198  
A:Molecule type: protein  
A:Residues: 565-575; 1890-1892, 'I', 1894-1900 <MAD>  
C:Comment: Fibrillin is a major component of elastin-associated microfibrils.  
C:Genetics:  
A:Gene: GDB:FBN1  
A:Cross-references: GDB:127115; OMIM:134797; OMIM:154700  
A:Map position: 15q21.1-15q21.1  
A:Introns: 2236/1; 2258/1; 2297/1  
C:Superfamily: fibrillin; EGF homology  
C:Keywords: alternative splicing; calcium binding; extracellular matrix; glycoprotein;  
F:1-3002/Product: fibrillin (5'-region exon A splice form) #status predicted <MATC>  
F:1332-3002/Product: fibrillin (5'-region exon C splice form) #status predicted <MATC>  
F:1332-1367/Domain: EGF homology <EGF>  
F:1457-1492/Domain: EGF homology <EGF2>  
F:2262-2295/Domain: EGF homology <EGF1>

Query Match 11.7%; Score 176.5; DB 2; Length 3002;  
Best Local Similarity 22.7%; Pred. No. 6e-05;  
Matches 59; Conservative 19; Mismatches 79; Indels 103; Gaps 8;

QY 26 PGRVCAVRAHGDVPSSEFVQRYVQPLTTCDGHR---ACSTYRTYTYARRSPLGAPA 82  
Db 206 PGGNQCIVPICRHS CGDGFCSR---PNNCTCPSGQIAPSCGS-RSIQHCNIRCMNGGSCS 261  
QY 83 RPRYACCPGWKRTSLGACGAATCOPPCNGSGSCVQGRCPAGWGTGTCQ----- 135  
Db 262 DDHCLCKQKYGITH-----CQPVCSGCLNGRCVAPNRCACCTYGTGQCCEYATGP 316  
QY 136 -----S 136  
Db 317 CFTVISNQMCQQLSGIVCTKQLCCATVGRAWHGPCMCPAQHPHRRGFIPNIRTGACQ 376  
QY 137 DVDECSARRGGCP-ORCINTAGSYWCQWEGHSLSDGTLG-----VPKGG--- 181  
Db 377 DVDECSAIPGICQCGNCINTVGSFECCKPAGHKLNVSKQKEDIDECSTIPGICEGSECT 436  
QY 182 -----PPRVAPNPTG 191  
Db 437 NTVSSYFCKCPPGFTVSDG 455

RESULT 12

A>Title: The protein encoded by a growth arrest-specific gene (gas6) is a new member of

A:Reference number: A48089; MUID:93330291; PMID:8336730  
A:Accession: B48089  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-678 <MAN>  
A:Cross-references: GB:L13720; NID:G401766; PIDN:AAA58494.1; PID:G401767  
C:Superfamily: plasma protein S; EGF homology; Gla domain homology; laminin G repeat hom  
F:41-92/Domain: Gla domain homology #status atypical <GLA>  
F:120-153/Domain: EGF homology <EG1>  
F:160-195/Domain: EGF homology <EG2>  
F:201-236/Domain: EGF homology <EG3>  
F:242-277/Domain: EGF homology <EG4>  
F:311-671/Domain: sex hormone-binding globulin homology <SHB>  
F:321-473/Domain: laminin G repeat homology <LGR>

Query Match 12.6%; Score 189; DB 2; Length 678;  
Best Local Similarity 34.2%; Pred. No. 2e-06;  
Matches 53; Conservative 8; Mismatches 64; Indels 30; Gaps 8;

QY 46 QRYVQFFLTCDGH-----RACSTYRTIYRTYRSPGLAPRPRVACC-----PGWK 93

DB 47 RRAQVPEAKQGHLEECVEVCSKEA--REVFPNDPTEYFPRYDCINKYGRPEE 104

QY 94 RTSG-----LPGACGAAICQPPC--RNGSCVOP--GR--CRCPAGWRGDTQSDVDEC 141

DB 105 KNSGFATCVQNLPOC-----TFNFCDRKGTQACQDLGNFFCLCKAGWGRLCDKDVNEC 160

QY 142 SARGGCGPQRCINTAGSYWCQWEGHSLASDGTLC 176

DB 161 SQENGGLQICHNKPFGFCHSGFELSDGRTC 195

A>Title: The protein encoded by a growth arrest-specific gene (gas6) is a new member of

A:Reference number: A48089; MUID:93330291; PMID:8336730  
A:Accession: A48089  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-673 <MAN>  
A:Cross-references: GB:X59846; NID:G407060; PIDN:CAA42507.1; PID:G407061  
A>Note: authors translated the codon CCC for residue 424 as Ile  
C:Genetics:  
A:Gene: gas6  
C:Superfamily: plasma protein S; EGF homology; Gla domain homology; laminin G repeat hom  
F:38-89/Domain: Gla domain homology #status atypical <GLA>  
F:117-150/Domain: EGF homology <EG1>  
F:157-192/Domain: EGF homology <EG2>  
F:198-233/Domain: EGF homology <EG3>  
F:238-274/Domain: EGF homology <EG4>  
F:308-666/Domain: sex hormone-binding globulin homology <SHB>  
F:318-470/Domain: laminin G repeat homology <LGR>

Query Match 12.2%; Score 184; DB 2; Length 673;  
Best Local Similarity 31.6%; Pred. No. 4.6e-06;  
Matches 50; Conservative 8; Mismatches 64; Indels 36; Gaps 6;

QY 46 QRYVQFFLTCDGH-----RACSTYRTIYRTYRSPGLAPRPRVACC----- 89

DB 44 RRAYQVPEAKQGHLEECVEVCSKEA--REVFPNDPTEYFPRYDCINKYGRPEE 101

QY 90 --PGW-KETSLGPLAC-----GAAICQPPCPRNGSCVOPGRCRCPAGWRGDTQSDV 138

DB 102 KNFDFAKCVQNLPOCCTNFCDDKGTGHCQDLGN-----FFCVCTDGGWGLCDKDV 154

QY 139 DECSARGGCPQRCINTAGSYWCQWEGHSLASDGTLC 176

DB 155 NECVQKNGGCSQVCHNKPFGFCHSGFELSDGRTC 192

#### RESULT 8

A57278

fibrillin-2 precursor - mouse

C:Species: Mus musculus (house mouse)

C>Date: 23-Feb-1996 #sequence\_revision 23-Feb-1996 #text\_change 24-Nov-2003

C:Accession: A57278

R:Zhang, H.; Hu, W.; Ramirez, F.

J. Cell Biol. 129, 1165-1176, 1995

A>Title: Developmental expression of fibrillin genes suggests heterogeneity of extrace

A:Reference number: A57278; MUID:95263670; PMID:7744963

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-2907 <ZHA>

A:Cross-references: GB:L139790; NID:G762830; PIDN:AAA74908.1; PID:G762831

C:Superfamily: fibrillin; EGF homology

F:1239-1274/Domain: EGF homology <EGF1>

F:2488-2523/Domain: EGF homology <EGF>

Query Match 12.2%; Score 183.5; DB 2; Length 2907;

Best Local Similarity 23.9%; Pred. No. 1.8e-05;

Matches 60; Conservative 19; Mismatches 85; Indels 87; Gaps 7;

QY 2 RGQEVLLMWLLVAVGGTEHAY-----RGRRCVAVRAHGDVPVSEFVQVYQPPIT 54

DB 77 RGQEVLLMWLLVAVGGTEHAY-----RGRRCVAVRAHGDVPVSEFVQVYQPPIT 54

QY 55 TCDGHEACST--VRTIYRTYRSPGLAPRPRVACCPGKRTSLGPGAGAAICQPPCR 112

DB 131 TSSGQISPTGKRSIQCSVROMNGTCDHCCQKGIYGT-----YCGPVCENGQ 185

QY 113 NGSCVQPCRCPCAGWRGDTQ-----

DB 186 NGRCIGPNCACVYGTGTCQPCDRTYGTCTVNNMCQQLTGIVCTKTLCCATIGR 245

QY 136 -----SDYDECSARGGCP-ORCINTAGSYWCQW 165

DB 246 ANGHPC 305

QY 166 GHSLASDGTLC 176

DB 306 GHKQSETTQKC 316

A>Title: The protein encoded by a growth arrest-specific gene (gas6) is a new member of

A:Reference number: A48089; MUID:93330291; PMID:8336730

A:Accession: A48089

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-673 <MAN>

A:Cross-references: GB:X59846; NID:G407060; PIDN:CAA42507.1; PID:G407061

A>Note: authors translated the codon CCC for residue 424 as Ile

C:Genetics:  
A:Gene: gas6  
C:Superfamily: plasma protein S; EGF homology; Gla domain homology; laminin G repeat hom

F:38-89/Domain: Gla domain homology #status atypical <GLA>

F:117-150/Domain: EGF homology <EG1>

F:157-192/Domain: EGF homology <EG2>

F:198-233/Domain: EGF homology <EG3>

F:238-274/Domain: EGF homology <EG4>

F:308-666/Domain: sex hormone-binding globulin homology <SHB>

F:318-470/Domain: laminin G repeat homology <LGR>

Query Match 12.2%; Score 184; DB 2; Length 673;  
Best Local Similarity 31.6%; Pred. No. 4.6e-06;  
Matches 50; Conservative 8; Mismatches 64; Indels 36; Gaps 6;

QY 46 QRYVQFFLTCDGH-----RACSTYRTIYRTYRSPGLAPRPRVACC----- 89

DB 44 RRAYQVPEAKQGHLEECVEVCSKEA--REVFPNDPTEYFPRYDCINKYGRPEE 101

QY 90 --PGW-KETSLGPLAC-----GAAICQPPCPRNGSCVOPGRCRCPAGWRGDTQSDV 138

DB 102 KNFDFAKCVQNLPOCCTNFCDDKGTGHCQDLGN-----FFCVCTDGGWGLCDKDV 154

#### RESULT 9

T10756

Nel-homolog protein - rat

C:Species: Rattus norvegicus (Norway rat)

C>Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 16-Jul-1999

C:Accession: T10756

R:Kuroda, S.; Tokunaga, C.; Kiyohara, Y.; Konishi, H.; Matsuhashi, S.; Kikkawa, U.

submitted to the EMBL Data Library, November 1998

A:Description: Protein kinase C-binding protein.

A:Reference number: Z17122

A:Accession: T10756

A>Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: mRNA

A:Residues: 1-810 <KUR>

A:Cross-references: EMBL:U48246; NID:G3851179; PID:G3851180

A:Experimental source: strain Sprague-Dawley, brain

Query Match 12.0%; Score 180; DB 2; Length 810;

Best Local Similarity 36.5%; Pred. No. 1.1e-05;

Matches 35; Conservative 13; Mismatches 38; Indels 10; Gaps 3;

QY 88 CCGGKRTSLGPGACAAICQPPCPRNGSCVOPGRCRCPAGWRGDTQSDVDECSARRGG 147

R; Nakayama, M.; Nakajima, D.; Nagase, T.; Nomura, N.; Seki, N.; Ohara, O.  
Genomics 51, 27-34, 1998  
A; Title: Identification of high-molecular-weight proteins with multiple EGF-like motifs  
A; Reference number: Z14126; MUID: 98360089; PMID: 9693030  
A; Accession: T13954  
A; Status: preliminary; translated from GB/EMBL/DBJ  
A; Molecule type: mRNA  
A; Residues: 1-1574 <NAK>  
A; Cross-references: EMBL:AB011532; NID: g3449293; PIDN: BAA32462.1; PID: g3449294  
A; Experimental source: strain Sprague-Dawley; brain  
C; Genetics:  
A; Gene: MEGP6

Query Match 17.0%; Score 256; DB 2; Length 1574;  
Best Local Similarity 37.9%; Pred. No. 5.2e-11;  
Matches 64; Conservative 16; Mismatches 61; Indels 28; Gaps 9;  
QY 30 VCAVR-----AHGDPVSEGFVQVRYVQPLTTCDGHRAC--STYRTYRTAYRSPGLAPA 82  
DB 43 VCAEQKLTLVGHRQPCVQAFSRIVPVVRRITGCAQQAQWCIQGERITVYMSYQVYA-TEA 101  
QY 83 RPRYACCPGWKRTSGLPGA-----CGAAL--CQPPCN--CGSCVQPGRCPCAGWR-- 130  
DB 102 RTVRCPCGWSQKPGQEGCLSDVDECAANGCGEGCCNTVGGF-----YCRCPFGYQLQ 156  
QY 131 --GDTCCSDVDECSARRGGPCPCINTAGSYWCQWEGHSLGADGTLV 177  
DB 157 GDGKTIQ-DVDECRANGGCGQRCVNTPGSYICEKPGFLATDGTCL 204

## RESULT 3

T17324  
hypothetical protein DKFP564P2063.1 - human (fragment)  
C; Species: Homo sapiens (man)  
C; Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999  
C; Accession: T17324  
R; Duesterhoeft, A.; Lauber, J.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.  
submitted to the Protein Sequence Database, September 1999  
A; Reference number: Z18727  
A; Accession: T17324  
A; Status: preliminary  
A; Molecule type: mRNA  
A; Residues: 1-558 <DUE>  
A; Cross-references: EMBL: A1117610  
A; Experimental source: fetal brain; clone DKFP564P2063  
C; Genetics:  
A; Note: DKFP564P2063.1

Query Match 16.8%; Score 252.5; DB 2; Length 558;  
Best Local Similarity 36.8%; Pred. No. 3.8e-11;  
Matches 57; Conservative 17; Mismatches 50; Indels 31; Gaps 7;  
QY 58 GHRACSTYRTYRTAYRSPGLAPAPRVYACCPGWKRTSGLPGACGAAICQPPCRNGGSC 117  
DB 26 GNAASARHHGLASA--RQGVCHYGTKLACCYGWRRNS--KGVCEATCEPGCK-FGEC 79  
QY 118 VQGRRCPCAGWGDTCQSDVDECSARRGGPCPCPCINTAGSYWCQWEGHSLGADGTLV 177  
DB 80 VGNPKRCCLPGYTGKTCQSDVDECSARRGGPCPCPCINTAGSYWCQWEGHSLGADGTLV 177  
QY 178 -----PKGFGPP-----RVAPN 188  
DB 139 YSRTCAMINQYSCDTEEGPQCLCPSSGLRLAPN 173

## RESULT 4

T27283  
hypothetical protein Y64G10A.f - Caenorhabditis elegans  
C; Species: Caenorhabditis elegans  
C; Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999  
C; Accession: T27283  
R; Ainscough, R.  
submitted to the EMBL Data Library, September 1999

A; Reference number: Z20336  
A; Accession: T27283  
A; Status: preliminary; translated from GB/EMBL/DBJ  
A; Molecule type: DNA  
A; Residues: 1-1620 <WIL>  
A; Cross-references: EMBL: A1110498; NID: s1542303; PIDN: CAB54471.1; CESP: Y64G10A.f  
A; Experimental source: clone Y64G10A  
C; Genetics:  
A; Gene: CESP: Y64G10A.f  
A; Introns: 77/1; 116/1; 198/1; 282/1; 365/1; 425/1; 466/1; 548/1; 559/1; 601/1; 625/1;  
Query Match 14.5%; Score 218; DB 2; Length 1620;  
Best Local Similarity 40.4%; Pred. No. 3.2e-08;  
Matches 42; Conservative 9; Mismatches 35; Indels 18; Gaps 4;  
QY 102 CGAIAICPPCRNGGSCVQPGR-----CRCPAGWRGTCQSDVDECSARRGGPCPCINT 155  
DB 79 CSADL-----CHNGGTCTVPSEHNDNEQVCECFVGTGAKCOYDANECWANNNGGCEHCVT 134  
QY 156 AGSYWCQWEGHSLGADGTLV-----VFKGG-PPRVAPNPTG 191  
DB 135 IGTYICRWFQFELSGDGTCSIDECVANGGSDRCVNSPG 178

## RESULT 5

T09059  
notch4 - mouse  
C; Species: Mus musculus (house mouse)  
C; Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 08-Sep-2002  
C; Accession: T09059  
R; Rowen, L.; Mahairas, G.; Qin, S.; Ahearn, M.E.; Dankers, C.; Loretz, C.;  
submitted to the EMBL Data Library, October 1997  
A; Description: Sequence of the mouse major histocompatibility locus class III region.  
A; Reference number: Z16543  
A; Accession: T09059  
A; Status: preliminary; translated from GB/EMBL/DBJ  
A; Molecule type: DNA  
A; Residues: 1-1964 <ROW>  
A; Cross-references: EMBL: AF030001; NID: g2564945; PID: g2564947  
C; Genetics:  
A; Gene: notch4  
A; Map position: 17  
A; Introns: 22/1; 49/2; 148/1; 264/1; 305/1; 384/1; 436/1; 501/1; 539/1; 577/1; 618/1;  
1679/3; 1729/1; 1761/3  
C; Superfamily: notch protein; ankyrin repeat homology; EGF homology  
C; Keywords: receptor; signal transduction  
F; 514-545/Domain: EGF homology <EGF>

Query Match 13.0%; Score 196; DB 2; Length 1964;  
Best Local Similarity 34.2%; Pred. No. 1.6e-06;  
Matches 50; Conservative 12; Mismatches 54; Indels 30; Gaps 7;  
QY 80 APARPRYACCPGWK-RTSGLPGACGAAICQPPCRNGGSCVQPG---RCRCPCAGWRGTCQ 135  
DB 134 ASGRPQSCCEPGWTEGQCQLRDFCSA---NPCANGVCLATYPIQICRCPGPGFEGTCE 189  
QY 136 SDVDECSARRGGCQ--RCINTAGSYWCQ---WEGHSLGADGTLVCP-----KGGPRVA 186  
DB 190 RDINECFLEPGPCPGTGSCHNTLSYQCLCPVQGGPQCKLRKACPPGSLNGTCLQV 249  
QY 187 FNP-----TGVDSAMKEE 199  
DB 250 PEGHSTFHLCLCPPGFTGLDCMNP 275

## RESULT 6

B48089  
growth arrest-specific protein gas6 - human  
C; Species: Homo sapiens (man)  
C; Date: 26-May-1994 #sequence\_revision 26-May-1994 #text\_change 05-Nov-1999  
C; Accession: B48089  
R; Manfioletti, G.; Brancolini, C.; Avanzi, G.; Schneider, C.  
Mol. Cell. Biol. 13, 4976-4985, 1993







Query Match 14.4%; Score 216.5; DB 4; Length 537;  
Best Local Similarity 39.3%; Pred. No. 1.7e-09;  
Matches 48; Conservative 10; Mismatches 35; Indels 29; Gaps 6;  
QY 91 GWKRTSGLPGACGAATCQPPCRNGGSCVOPGRCRCPCAGWRGDTCCSDVDVDECSARRGGCPQ 150  
DB 1 GWRNS--KGVC-EATCEPGCK-FGECVGNKRCRCFPFGYTGTCTCSQDVNECGMKPRPCQH 56  
QY 151 RCINTAGSYWCQWEGHSLSDGTLVCV-----PKGGPP-----RVA 186  
DB 57 RCVNTGHSYKCFCLSGHMLMPDAT-CVNSRTCAMINCOYSCDTEEGPQCICLPSSGLRLA 115  
QY 187 PN 188  
DB 116 PN 117

RESULT 13  
US-09-249-697A-3  
; Sequence 3, Application US/09249697A  
; Patent No. 6392018  
; GENERAL INFORMATION:  
; APPLICANT: Ford, John  
; APPLICANT: Yeung, George  
; TITLE OF INVENTION: NOVEL EGF MOTIF PROTEIN OBTAINED FROM A CDNA LIBRARY OF FETAL  
; FILE REFERENCE: 24011-727  
; CURRENT APPLICATION NUMBER: US/09/249,697A  
; PRIOR FILING DATE: 1999-02-12  
; PRIOR FILING DATE: 1997-11-22  
; NUMBER OF SEQ ID NOS: 19  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 3  
; LENGTH: 100  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-249-697A-3

Query Match 14.3%; Score 215.5; DB 4; Length 100;  
Best Local Similarity 48.3%; Pred. No. 3.1e-10;  
Matches 42; Conservative 8; Mismatches 32; Indels 5; Gaps 4;  
QY 91 GWKRTSGLPGACGAATCQPPCRNGGSCVOPGRCRCPCAGWRGDTCCSDVDVDECSARRGGCPQ 150  
DB 1 GWRNS--KGVC-EATCEPGCK-FGECVGNKRCRCFPFGYTGTCTCSQDVNECGMKPRPCQH 56  
QY 151 RCINTAGSYWCQWEGHSLSDGTLVCV 177  
DB 57 RCVNTGHSYKCFCLSGHMLMPDAT-CV 82

RESULT 14  
US-09-363-316B-3  
; Sequence 3, Application US/09363316B  
; Patent No. 6392019  
; GENERAL INFORMATION:  
; APPLICANT: Ford, John  
; APPLICANT: Yeung, George  
; TITLE OF INVENTION: EGF MOTIF PROTEIN MATERIALS AND METHODS  
; FILE REFERENCE: 28110/35852  
; CURRENT APPLICATION NUMBER: US/09/363,316B  
; PRIOR FILING DATE: 1999-07-28  
; PRIOR FILING DATE: 1999-02-12  
; PRIOR FILING DATE: 1997-11-22  
; NUMBER OF SEQ ID NOS: 24  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 3  
; LENGTH: 100  
; TYPE: PRT

; ORGANISM: Homo sapiens  
US-09-363-316B-3  
Query Match 14.3%; Score 215.5; DB 4; Length 100;  
Best Local Similarity 48.3%; Pred. No. 3.1e-10;  
Matches 42; Conservative 8; Mismatches 32; Indels 5; Gaps 4;  
QY 91 GWKRTSGLPGACGAATCQPPCRNGGSCVOPGRCRCPCAGWRGDTCCSDVDVDECSARRGGCPQ 150  
DB 1 GWRNS--KGVC-EATCEPGCK-FGECVGNKRCRCFPFGYTGTCTCSQDVNECGMKPRPCQH 56  
QY 151 RCINTAGSYWCQWEGHSLSDGTLVCV 177  
DB 57 RCVNTGHSYKCFCLSGHMLMPDAT-CV 82

RESULT 15  
US-09-467-997-1  
; Sequence 1, Application US/09467997  
; Patent No. 6379925  
; GENERAL INFORMATION:  
; APPLICANT: Kitajewski, Jan  
; APPLICANT: Uyttendaele, Hendrik  
; TITLE OF INVENTION: ANGIOGENIC MODULATION BY NOTCH SIGNAL TRANSDUCTION  
; FILE REFERENCE: 53863-A-PCT-US  
; CURRENT APPLICATION NUMBER: US/09/467,997  
; CURRENT FILING DATE: 1999-12-20  
; NUMBER OF SEQ ID NOS: 10  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 1964  
; TYPE: PRT  
; ORGANISM: mouse  
US-09-467-997-1

Query Match 13.8%; Score 196; DB 4; Length 1964;  
Best Local Similarity 34.2%; Pred. No. 3.2e-07;  
Matches 50; Conservative 12; Mismatches 54; Indels 30; Gaps 7;  
QY 80 APARPRVACPCGK-RTSGLPGACGAATCQPPCRNGGSCVOPG---RCRCPCAGWRGDTCC 135  
DB 134 ASGRPQSCSEFGTGTGECQLRDFCSA-----NPCANGGVCLATYPQICRCFPFGHTCE 189  
QY 136 SDVDECSARRGGCPQ--RCINTAGSYWCQ---WEHSLSDGTLVCV-----KGGPRVA 186  
DB 190 RDINECFLEBPQCPQGTSCNTLGSYQCLCFVQGEQPCQKLRKGCPCPGSLNGGTCLV 249  
QY 187 PNP-----TGVDSAMKEE 199  
DB 250 PEGHSTPHLCLCPPGFTGLDCENPD 275

Search completed: April 22, 2004, 12:47:57  
Job time : 23 secs

```

; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 315
; LENGTH: 509
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-902-775A-315

Query Match      14.4%; Score 223; DB 4; Length 509;
Best Local Similarity 42.4%; Pred. No. 5e-10; Indels 4; Gaps 3;
Matches 42; Conservative 13; Mismatches 40;

QY 76 SPGLAPAPRYACCPGKRTSGLPCAGAAICQPPCRNGSCVQPGRCRCPCAGWRGDTQ 135
DB 32 SIGLCRYGGRIDCCGWARQSW--GQC-QPVCQPRKH-GEICIGPKCKCHPGYAGKTGN 87
QY 136 SDVDECSARRGGCQRQCRINTAGSYWCQCWEHSLSDGT 174
DB 88 QDLNECGLKPRFCKHRNMNTYGSYKCYCLNGYMLMPDGS 126

RESULT 10
US-09-363-316B-18
; Sequence 18, Application US/09363316B
; Patent No. 6392019
; GENERAL INFORMATION:
; APPLICANT: Ford, John
; TITLE OF INVENTION: EGF MOTIF PROTEIN MATERIALS AND METHODS
; FILE REFERENCE: 28110/35852
; CURRENT APPLICATION NUMBER: US/09/363,316B
; CURRENT FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: US 09/249,697
; PRIOR FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 08/968,800
; PRIOR FILING DATE: 1997-11-22
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 18
; LENGTH: 502
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: misc feature
; LOCATION: (501-502)
; OTHER INFORMATION: Xaa = any amino acid
US-09-363-316B-18

Query Match      14.4%; Score 216.5; DB 4; Length 502;
Best Local Similarity 39.3%; Pred. No. 1.6e-09;
Matches 48; Conservative 10; Mismatches 35; Indels 29; Gaps 6;

QY 91 GWKRTSGLPGACGAAICQPPCRNGSCVQPGRCRCPCAGWRGDTQSDVDECSARRGGC 150
DB 1 GWRNRS--KGVC-EATCEFGCK-FGECVGNKRCFPFGYTGKTCSDQDVNECGMKPRPC 56
QY 151 RCINTAGSYWCQCWEHSLSDGTLCV-----PKGPP-----RVA 186
DB 57 RCWNTHGSYKCFCLSGHMLMPDAT-CVNSRTCAMINCOYSCDTEBGPQCLCPSSGLRLA 115
QY 187 PN 188

RESULT 12
US-09-363-316B-4
; Sequence 4, Application US/09363316B
; Patent No. 6392019
; GENERAL INFORMATION:
; APPLICANT: Ford, John
; TITLE OF INVENTION: EGF MOTIF PROTEIN MATERIALS AND METHODS
; FILE REFERENCE: 28110/35852
; CURRENT APPLICATION NUMBER: US/09/363,316B
; CURRENT FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: US 09/249,697
; PRIOR FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 08/968,800
; PRIOR FILING DATE: 1997-11-22
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 537
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: misc feature
; LOCATION: (503)
; OTHER INFORMATION: Xaa = any amino acid
US-09-363-316B-4
```

```

DB 116 PN 117

RESULT 11
US-09-249-697A-4
; Sequence 4, Application US/09249697A
; Patent No. 6392018
; GENERAL INFORMATION:
; APPLICANT: Ford, John
; TITLE OF INVENTION: NOVEL EGF MOTIF PROTEIN OBTAINED FROM A CDNA LIBRARY OF FETAL
; TITLE OF INVENTION: LIVER SPLEEN
; FILE REFERENCE: 24011-727
; CURRENT APPLICATION NUMBER: US/09/249,697A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 08/968,800
; PRIOR FILING DATE: 1997-11-22
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 537
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: VARIANT
; LOCATION: (1)...(537)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-09-249-697A-4

Query Match      14.4%; Score 216.5; DB 4; Length 537;
Best Local Similarity 39.3%; Pred. No. 1.7e-09;
Matches 48; Conservative 10; Mismatches 35; Indels 29; Gaps 6;

QY 91 GWKRTSGLPGACGAAICQPPCRNGSCVQPGRCRCPCAGWRGDTQSDVDECSARRGGC 150
DB 1 GWRNRS--KGVC-EATCEFGCK-FGECVGNKRCFPFGYTGKTCSDQDVNECGMKPRPC 56
QY 151 RCINTAGSYWCQCWEHSLSDGTLCV-----PKGPP-----RVA 186
DB 57 RCWNTHGSYKCFCLSGHMLMPDAT-CVNSRTCAMINCOYSCDTEBGPQCLCPSSGLRLA 115
QY 187 PN 188

RESULT 12
US-09-363-316B-4
; Sequence 4, Application US/09363316B
; Patent No. 6392019
; GENERAL INFORMATION:
; APPLICANT: Ford, John
; TITLE OF INVENTION: EGF MOTIF PROTEIN MATERIALS AND METHODS
; FILE REFERENCE: 28110/35852
; CURRENT APPLICATION NUMBER: US/09/363,316B
; CURRENT FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: US 09/249,697
; PRIOR FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 08/968,800
; PRIOR FILING DATE: 1997-11-22
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 537
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: misc feature
; LOCATION: (503)
; OTHER INFORMATION: Xaa = any amino acid
US-09-363-316B-4
```

## RESULT 8

US-09-905-125A-315  
; Sequence 315, Application US/09905125A  
; Patent No. 6664376  
; GENERAL INFORMATION:  
; APPLICANT: Genentech, Inc.  
; APPLICANT: Ashkenazi, Avi  
; APPLICANT: Botstein, David  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Eaton, Dan L.  
; APPLICANT: Ferrara, Napoleone  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Fong, Sherman  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Gerber, Hanspeter  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, A.  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Grimaldi, Christopher J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Hillan, Kenneth, J.  
; APPLICANT: Kijavini, Ivar J.  
; APPLICANT: Mather, Jennie P.  
; APPLICANT: Pan, James  
; APPLICANT: Paoni, Nicholas F.  
; APPLICANT: Roy, Margaret Ann  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Williams, P. Mickey  
; APPLICANT: Wood, William, I.  
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
; TITLE OF INVENTION: Acids Encoding the Same  
; FILE REFERENCE: 10466-14  
; CURRENT APPLICATION NUMBER: US/09/905,125A  
; PRIOR FILING DATE: 2001-07-12  
; PRIOR APPLICATION NUMBER: PCT/US00/04414  
; PRIOR FILING DATE: 2000-02-22  
; PRIOR FILING DATE: 1999-07-07  
; PRIOR APPLICATION NUMBER: US 60/143,048  
; PRIOR FILING DATE: 1999-07-26  
; PRIOR APPLICATION NUMBER: US 60/145,698  
; PRIOR FILING DATE: 1999-07-28  
; PRIOR APPLICATION NUMBER: PCT/US99/20594  
; PRIOR FILING DATE: 1999-09-08  
; PRIOR APPLICATION NUMBER: PCT/US99/20944  
; PRIOR FILING DATE: 1999-09-13  
; PRIOR APPLICATION NUMBER: PCT/US99/21090  
; PRIOR FILING DATE: 1999-09-15  
; PRIOR APPLICATION NUMBER: PCT/US99/21547  
; PRIOR FILING DATE: 1999-09-15  
; PRIOR APPLICATION NUMBER: PCT/US99/23089  
; PRIOR FILING DATE: 1999-10-05  
; PRIOR APPLICATION NUMBER: PCT/US99/28214  
; PRIOR FILING DATE: 1999-11-29  
; PRIOR APPLICATION NUMBER: PCT/US99/28313  
; PRIOR FILING DATE: 1999-11-30  
; PRIOR APPLICATION NUMBER: PCT/US99/28564  
; PRIOR FILING DATE: 1999-12-02  
; PRIOR APPLICATION NUMBER: PCT/US99/28565  
; PRIOR FILING DATE: 1999-12-02  
; PRIOR APPLICATION NUMBER: PCT/US99/30095  
; PRIOR FILING DATE: 1999-12-16  
; PRIOR APPLICATION NUMBER: PCT/US99/30911  
; PRIOR FILING DATE: 1999-12-20  
; PRIOR APPLICATION NUMBER: PCT/US99/30999  
; PRIOR FILING DATE: 1999-12-20  
; PRIOR APPLICATION NUMBER: PCT/US00/00219  
; PRIOR FILING DATE: 2000-01-05  
; NUMBER OF SEQ ID NOS: 423  
; SEQ ID NO 315  
; LENGTH: 509

; TYPE: PRT  
; ORGANISM: Homo Sapien  
US-09-905-125A-315

Query Match 14.8%; Score 223; DB 4; Length 509;  
Best Local Similarity 42.4%; Pred. No. 5e-10;  
Matches 42; Conservative 13; Mismatches 40; Indels 4; Gaps 3;

QY	76	SPGLAPARPRYACCPGKTSGLPGAGAAICOPPCRNCGSSCVOPGRCPCPAGWRGDTQ	135
Db	32	SIGLCRYGRIIDCCWGAWQSW--GQC-OPVQPRCKH-GEICGNKCKHPGYAKTKCN	87
QY	136	SDVDECSARRGGPQRCINTAGSYWCQCWEGHSLSDGT	174
Db	88	QDLNECGLKPRCKHRCMNITYGSYKCYCLNGYMLMPDGS	126

## RESULT 9

US-09-902-775A-315  
; Sequence 315, Application US/09902775A  
; Patent No. 6686451  
; GENERAL INFORMATION:  
; APPLICANT: Genentech, Inc.  
; APPLICANT: Ashkenazi, Avi  
; APPLICANT: Botstein, David  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Eaton, Dan L.  
; APPLICANT: Ferrara, Napoleone  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Fong, Sherman  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Gerber, Hanspeter  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, A.  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Grimaldi, Christopher J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Hillan, Kenneth, J.  
; APPLICANT: Kijavini, Ivar J.  
; APPLICANT: Mather, Jennie P.  
; APPLICANT: Pan, James  
; APPLICANT: Paoni, Nicholas F.  
; APPLICANT: Roy, Margaret Ann  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Williams, P. Mickey  
; APPLICANT: Wood, William, I.  
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
; TITLE OF INVENTION: Acids Encoding the Same  
; FILE REFERENCE: 10466-14  
; CURRENT APPLICATION NUMBER: US/09/902,775A  
; PRIOR FILING DATE: 2001-07-10  
; PRIOR APPLICATION NUMBER: PCT/US00/04414  
; PRIOR FILING DATE: 2000-02-22  
; PRIOR APPLICATION NUMBER: US 60/143,048  
; PRIOR FILING DATE: 1999-07-07  
; PRIOR APPLICATION NUMBER: US 60/145,698  
; PRIOR FILING DATE: 1999-07-26  
; PRIOR APPLICATION NUMBER: PCT/US99/20594  
; PRIOR FILING DATE: 1999-07-28  
; PRIOR APPLICATION NUMBER: PCT/US99/20944  
; PRIOR FILING DATE: 1999-09-13  
; PRIOR APPLICATION NUMBER: PCT/US99/21090  
; PRIOR FILING DATE: 1999-09-15  
; PRIOR APPLICATION NUMBER: PCT/US99/21547  
; PRIOR FILING DATE: 1999-09-15  
; PRIOR APPLICATION NUMBER: PCT/US99/23089  
; PRIOR FILING DATE: 1999-10-05  
; PRIOR APPLICATION NUMBER: PCT/US99/28214  
; PRIOR FILING DATE: 1999-11-29  
; PRIOR APPLICATION NUMBER: PCT/US99/28313  
; PRIOR FILING DATE: 1999-11-30  
; PRIOR APPLICATION NUMBER: PCT/US99/28564  
; PRIOR FILING DATE: 1999-12-02  
; PRIOR APPLICATION NUMBER: PCT/US99/28565  
; PRIOR FILING DATE: 1999-12-02  
; PRIOR APPLICATION NUMBER: PCT/US99/30095  
; PRIOR FILING DATE: 1999-12-16  
; PRIOR APPLICATION NUMBER: PCT/US99/30911  
; PRIOR FILING DATE: 1999-12-20  
; PRIOR APPLICATION NUMBER: PCT/US99/30999  
; PRIOR FILING DATE: 1999-12-20  
; PRIOR APPLICATION NUMBER: PCT/US00/00219  
; PRIOR FILING DATE: 2000-01-05  
; NUMBER OF SEQ ID NOS: 423  
; SEQ ID NO 315  
; LENGTH: 509

Query Match 16.7%; Score 251.5; DB 4; Length 553;  
Best Local Similarity 36.8%; Pred. No. 2.9e-12;  
Matches 57; Conservative 17; Mismatches 50; Indels 31; Gaps 7;  
QY 58 GHRACSTYRTIYRARRSPGLAPRPRYACCPGKRTSLPAGCGAAICQPPCRNGGSC 117  
DB 21 GNAASARHGLLASA--RQPGVCHYGTKLACCYGWRNS--KGVC-EATCEPGCK-FGEC 74  
QY 118 VQGRCRCPAGWGTGTCSDVDECSARRGGCPQRCINTAGSYWCQCWEGHSLSDGTLV 177  
DB 75 VGNKRCRCPFGYTGKTCSDVNECGMKRPPCQHRVCVNTGSKYKFCFLSGHMLMPDAT-CV 133  
QY 178 -----PKGGPP-----RVAPN 188  
DB 134 NSRTCAMINCOYSCEDTEEGPQCLCPSSGLRLAPN 168

## RESULT 6

US-09-363-316B-24  
; Sequence 24, Application US/09363316B  
; Patent No. 6392019  
; GENERAL INFORMATION:  
; APPLICANT: Ford, John  
; APPLICANT: Yeung, George  
; TITLE OF INVENTION: EGF MOTIF PROTEIN MATERIALS AND METHODS  
; FILE REFERENCE: 28110/35852  
; CURRENT APPLICATION NUMBER: US/09/363.316B  
; CURRENT FILING DATE: 1999-07-28  
; PRIOR APPLICATION NUMBER: US 09/249,697  
; PRIOR FILING DATE: 1999-02-12  
; PRIOR APPLICATION NUMBER: US 08/968,800  
; PRIOR FILING DATE: 1997-11-22  
; NUMBER OF SEQ ID NOS: 24  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 24  
; LENGTH: 553  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-363-316B-24

Query Match 16.7%; Score 251.5; DB 4; Length 553;  
Best Local Similarity 36.8%; Pred. No. 2.9e-12;  
Matches 57; Conservative 17; Mismatches 50; Indels 31; Gaps 7;  
QY 58 GHRACSTYRTIYRARRSPGLAPRPRYACCPGKRTSLPAGCGAAICQPPCRNGGSC 117  
DB 21 GNAASARHGLLASA--RQPGVCHYGTKLACCYGWRNS--KGVC-EATCEPGCK-FGEC 74  
QY 118 VQGRCRCPAGWGTGTCSDVDECSARRGGCPQRCINTAGSYWCQCWEGHSLSDGTLV 177  
DB 75 VGNKRCRCPFGYTGKTCSDVNECGMKRPPCQHRVCVNTGSKYKFCFLSGHMLMPDAT-CV 133  
QY 178 -----PKGGPP-----RVAPN 188  
DB 134 NSRTCAMINCOYSCEDTEEGPQCLCPSSGLRLAPN 168

## RESULT 7

US-09-907-794A-315  
; Sequence 315, Application US/09907794A  
; Patent No. 6635468  
; GENERAL INFORMATION:  
; APPLICANT: Genentech, Inc.  
; APPLICANT: Ashkenazi, Avi  
; APPLICANT: Botstein, David  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Eaton, Dan L.  
; APPLICANT: Ferrara, Napoleone  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Fong, Sherman  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Gerber, Hanspeter

; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, A.  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Grimaldi, Christopher J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Hillan, Kenneth, J.  
; APPLICANT: Kljavin, Ivar J.  
; APPLICANT: Mather, Jennie P.  
; APPLICANT: Pan, James  
; APPLICANT: Paoni, Nicholas F.  
; APPLICANT: Roy, Margaret Ann.  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Williams, P. Mickey  
; APPLICANT: Wood, William, I.  
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
; TITLE OF INVENTION: Acids Encoding the Same  
; FILE REFERENCE: 10466-14  
; CURRENT APPLICATION NUMBER: US/09/907,794A  
; CURRENT FILING DATE: 2001-07-17  
; PRIOR APPLICATION NUMBER: PCT/US00/04414  
; PRIOR FILING DATE: 2000-02-22  
; PRIOR APPLICATION NUMBER: US 60/143,048  
; PRIOR FILING DATE: 1999-07-07  
; PRIOR APPLICATION NUMBER: US 60/145,698  
; PRIOR FILING DATE: 1999-07-26  
; PRIOR APPLICATION NUMBER: US 60/146,222  
; PRIOR FILING DATE: 1999-07-28  
; PRIOR APPLICATION NUMBER: PCT/US99/20594  
; PRIOR FILING DATE: 1999-09-08  
; PRIOR APPLICATION NUMBER: PCT/US99/20944  
; PRIOR FILING DATE: 1999-09-13  
; PRIOR APPLICATION NUMBER: PCT/US99/21090  
; PRIOR FILING DATE: 1999-09-15  
; PRIOR APPLICATION NUMBER: PCT/US99/21547  
; PRIOR FILING DATE: 1999-09-15  
; PRIOR APPLICATION NUMBER: PCT/US99/23089  
; PRIOR FILING DATE: 1999-10-05  
; PRIOR APPLICATION NUMBER: PCT/US99/28214  
; PRIOR FILING DATE: 1999-11-29  
; PRIOR APPLICATION NUMBER: PCT/US99/28313  
; PRIOR FILING DATE: 1999-11-30  
; PRIOR APPLICATION NUMBER: PCT/US99/28564  
; PRIOR FILING DATE: 1999-12-02  
; PRIOR APPLICATION NUMBER: PCT/US99/28565  
; PRIOR FILING DATE: 1999-12-02  
; PRIOR APPLICATION NUMBER: PCT/US99/30095  
; PRIOR FILING DATE: 1999-12-16  
; PRIOR APPLICATION NUMBER: PCT/US99/30911  
; PRIOR FILING DATE: 1999-12-20  
; PRIOR APPLICATION NUMBER: PCT/US99/30999  
; PRIOR FILING DATE: 1999-12-20  
; PRIOR APPLICATION NUMBER: PCT/US00/00219  
; PRIOR FILING DATE: 2000-01-05  
; NUMBER OF SEQ ID NOS: 423  
; SEQ ID NO 315  
; LENGTH: 509  
; TYPE: PRT  
; ORGANISM: Homo Sapien  
US-09-907-794A-315

Query Match 14.8%; Score 223; DB 4; Length 509;  
Best Local Similarity 42.4%; Pred. No. 5e-10;  
Matches 42; Conservative 13; Mismatches 40; Indels 4; Gaps 3;  
QY 76 SPGLAPRPRYACCPGKRTSLPAGCGAAICQPPCRNGGSCVQPCRCPCAGWGTGTCQ 135  
DB 32 SIGLCRYGGRIDCCWGRQSW--GCC-QPVCQPCCKH-GECIGNKCKCHPGYAGKTGN 87  
QY 136 SDVDECSARRGGCPQRCINTAGSYWCQCWEGHSLSDGTLV 174  
DB 88 QDLNECGKLPKPCPKCHRCMNTYGSYKCYCLNGYMLMPDGS 126

Patent No. 6392018  
GENERAL INFORMATION:  
APPLICANT: Yeung, George  
TITLE OF INVENTION: NOVEL EGF MOTIF PROTEIN OBTAINED FROM A CDNA LIBRARY OF FETAL LIVER SPLEEN  
FILE REFERENCE: 24011-727  
CURRENT APPLICATION NUMBER: US/09/249,697A  
PRIOR FILING DATE: 1999-02-12  
PRIOR APPLICATION NUMBER: US 08/968,800  
PRIOR FILING DATE: 1997-11-22  
NUMBER OF SEQ ID NOS: 19  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 6  
LENGTH: 553  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: VARIANT  
LOCATION: (1)...(553)  
OTHER INFORMATION: Xaa = Any Amino Acid  
US-09-249-697A-6

Query Match 16.9%; Score 254.5; DB 4; Length 553;  
Best Local Similarity 34.5%; Pred. No. 1.7e-12;  
Matches 60; Conservative 21; Mismatches 62; Indels 31; Gaps 7;

QY 39 PVSESFVQVYQPFLLTCDDGHRACSTYRTIYTYRSPGLAPAPRYACCPGKRTSGL 98  
Db 2 PLPWSIALPLLVPWAGFGNNAARHGLLSA--RQGVCHYTKLACCVGWRNS-- 57

QY 99 PGACGAACQPPCRNGGSCVQPCRCPCAGWRGDTCCSDVDDECSARRGCGPCRCINTAGS 158  
Db 58 KGVC-EATCEPGCK-FGECVGNKCRCPFGYTGKTCSDVNECGMKPRPCQHRVCNTHGS 115

QY 159 YWCQWEGHSLADGTLV-----PKGGPP-----RVAPN 188  
Db 116 YKCFCLSGHMLMPDAT-CVNSRTAMINCQYSCDETEGFCQCLPSSGLRLAPN 168

Query Match 16.9%; Score 254.5; DB 4; Length 553;  
Best Local Similarity 34.5%; Pred. No. 1.7e-12;  
Matches 60; Conservative 21; Mismatches 62; Indels 31; Gaps 7;

QY 39 PVSESFVQVYQPFLLTCDDGHRACSTYRTIYTYRSPGLAPAPRYACCPGKRTSGL 98  
Db 2 PLPWSIALPLLVPWAGFGNNAARHGLLSA--RQGVCHYTKLACCVGWRNS-- 57

QY 99 PGACGAACQPPCRNGGSCVQPCRCPCAGWRGDTCCSDVDDECSARRGCGPCRCINTAGS 158  
Db 58 KGVC-EATCEPGCK-FGECVGNKCRCPFGYTGKTCSDVNECGMKPRPCQHRVCNTHGS 115

QY 159 YWCQWEGHSLADGTLV-----PKGGPP-----RVAPN 188  
Db 116 YKCFCLSGHMLMPDAT-CVNSRTAMINCQYSCDETEGFCQCLPSSGLRLAPN 168

Query Match 16.9%; Score 254.5; DB 4; Length 553;  
Best Local Similarity 34.5%; Pred. No. 1.7e-12;  
Matches 60; Conservative 21; Mismatches 62; Indels 31; Gaps 7;

QY 39 PVSESFVQVYQPFLLTCDDGHRACSTYRTIYTYRSPGLAPAPRYACCPGKRTSGL 98  
Db 2 PLPWSIALPLLVPWAGFGNNAARHGLLSA--RQGVCHYTKLACCVGWRNS-- 57

QY 99 PGACGAACQPPCRNGGSCVQPCRCPCAGWRGDTCCSDVDDECSARRGCGPCRCINTAGS 158  
Db 58 KGVC-EATCEPGCK-FGECVGNKCRCPFGYTGKTCSDVNECGMKPRPCQHRVCNTHGS 115

QY 159 YWCQWEGHSLADGTLV-----PKGGPP-----RVAPN 188  
Db 116 YKCFCLSGHMLMPDAT-CVNSRTAMINCQYSCDETEGFCQCLPSSGLRLAPN 168

Db 2 PLPWSIALPLLVPWAGFGNNAARHGLLSA--RQGVCHYTKLACCVGWRNS-- 57

QY 99 PGACGAACQPPCRNGGSCVQPCRCPCAGWRGDTCCSDVDDECSARRGCGPCRCINTAGS 158

Db 58 KGVC-EATCEPGCK-FGECVGNKCRCPFGYTGKTCSDVNECGMKPRPCQHRVCNTHGS 115

QY 159 YWCQWEGHSLADGTLV-----PKGGPP-----RVAPN 188

Db 116 YKCFCLSGHMLMPDAT-CVNSRTAMINCQYSCDETEGFCQCLPSSGLRLAPN 168

RESULT 4  
US-09-312-283C-389  
Sequence 382, Application US/09312283C  
Patent No. 6573095  
GENERAL INFORMATION:  
APPLICANT: Watson, James D.  
APPLICANT: Strachan, Lorna  
APPLICANT: Sleeman, Matthew  
APPLICANT: Onrust, Rene  
APPLICANT: Marison, James G.  
APPLICANT: Kumble, Krishanand D.  
TITLE OF INVENTION: Compositions Isolated from Skin Cells  
TITLE OF INVENTION: and Methods for Their Use  
FILE REFERENCE: 11000.1011G2  
CURRENT APPLICATION NUMBER: US/09/312,283C  
CURRENT FILING DATE: 1999-05-14  
NUMBER OF SEQ ID NOS: 425  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 389  
LENGTH: 284  
TYPE: PRT  
ORGANISM: Mouse  
US-09-312-283C-389

Query Match 16.8%; Score 252.5; DB 4; Length 284;  
Best Local Similarity 36.3%; Pred. No. 1.1e-12;  
Matches 53; Conservative 15; Mismatches 51; Indels 27; Gaps 4;

QY 66 RTIYTYRSPGLAPAPRYACCPGKRTSGLPAGACGAACQPPCRNGGSCVQPCRCRC 125  
Db 47 RDYWLPAHAHQPGVCHYTKTACCYGNKNS--KGVC-EAVCEPRCK-FGECVGNKCR 102

QY 126 PAGWRGDTCCSDVDDECSARRGCGPCRCINTAGSYWCQWEGHSL----- 169

Db 103 FPGYTGKTCSDVNECAFRPCQHRVCNTHGSYKCFCLSGHMLLPDATCSNRTCARIN 162

QY 170 -----SADGTLVCPKGGPPRVAPN 188

Db 163 CQYSCDETAGPRCVCPPSSGLRLGPN 188

RESULT 5  
US-09-249-697A-19  
Sequence 19, Application US/09249697A  
Patent No. 6392018  
GENERAL INFORMATION:  
APPLICANT: Ford, John  
APPLICANT: Yeung, George  
TITLE OF INVENTION: NOVEL EGF MOTIF PROTEIN OBTAINED FROM A CDNA LIBRARY OF FETAL LIVER SPLEEN  
FILE REFERENCE: 24011-727  
CURRENT APPLICATION NUMBER: US/09/249,697A  
CURRENT FILING DATE: 1999-02-12  
PRIOR APPLICATION NUMBER: US 08/968,800  
PRIOR FILING DATE: 1997-11-22  
NUMBER OF SEQ ID NOS: 19  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 19  
LENGTH: 553  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-249-697A-19

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 22, 2004, 12:44:11 ; Search time 22 Seconds  
(without alignments)  
640.631 Million cell updates/sec

Title: US-09-978-191a-506  
Perfect score: 1505  
Sequence: 1 MRGSEVLLMWLLVAVGGT.....SEQISFLEQLGSCSKKDS 273

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:\*  
1: /cgn2\_6/ptodata/2/iaa/5A COMB.pep.\*  
2: /cgn2\_6/ptodata/2/iaa/5B COMB.pep.\*  
3: /cgn2\_6/ptodata/2/iaa/6A COMB.pep.\*  
4: /cgn2\_6/ptodata/2/iaa/6B COMB.pep.\*  
5: /cgn2\_6/ptodata/2/iaa/PTUS COMB.pep.\*  
6: /cgn2\_6/ptodata/2/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1180	78.4	278	4	US-09-724-864-52
2	254.5	16.9	553	4	US-09-249-697A-6
3	254.5	16.9	553	4	US-09-363-316B-6
4	252.5	16.8	284	4	US-09-312-283C-389
5	251.5	16.7	553	4	US-09-249-697A-19
6	251.5	16.7	553	4	US-09-363-316B-24
7	223	14.8	509	4	US-09-907-794A-315
8	223	14.8	509	4	US-09-905-125A-315
9	223	14.8	509	4	US-09-902-775A-315
10	216.5	14.4	502	4	US-09-363-316B-18
11	216.5	14.4	537	4	US-09-249-697A-4
12	216.5	14.4	537	4	US-09-363-316B-4
13	215.5	14.3	100	4	US-09-249-697A-3
14	215.5	14.3	100	4	US-09-363-316B-3
15	196	13.0	1964	4	US-09-467-997-1
16	189	12.6	678	1	US-08-282-141-2
17	189	12.6	678	1	US-08-435-434-2
18	189	12.6	678	1	US-08-435-436-2
19	189	12.6	678	2	US-08-438-863-2
20	189	12.6	678	2	US-08-438-864-2
21	189	12.6	678	3	US-08-438-862-2
22	189	12.6	678	3	US-08-628-747-2
23	189	12.6	678	3	US-08-402-253-2
24	189	12.6	678	3	US-08-443-866B-2
25	184	12.2	673	1	US-08-282-141-3
26	184	12.2	673	1	US-08-435-434-1
27	184	12.2	673	1	US-08-435-436-1

28 184 12.2 673 2 US-08-438-863-1 Sequence 1, Appli  
29 184 12.2 673 2 US-08-438-864-1 Sequence 1, Appli  
30 184 12.2 673 3 US-08-438-862-1 Sequence 1, Appli  
31 184 12.2 673 3 US-08-628-747-1 Sequence 1, Appli  
32 184 12.2 673 3 US-08-402-253-1 Sequence 1, Appli  
33 184 12.2 673 3 US-08-443-866B-1 Sequence 1, Appli  
34 182.5 12.1 816 3 US-08-820-170A-37 Sequence 37, Appli  
35 182.5 12.1 816 3 US-09-055-699-37 Sequence 37, Appli  
36 182.5 12.1 816 3 US-09-273-565-37 Sequence 37, Appli  
37 182.5 12.1 816 4 US-09-565-538-37 Sequence 37, Appli  
38 182.5 12.1 816 4 US-09-661-468-37 Sequence 37, Appli  
39 182.5 12.1 816 4 US-09-976-165-37 Sequence 37, Appli  
40 182 12.1 810 2 US-08-820-170A-34 Sequence 34, Appli  
41 182 12.1 810 3 US-09-055-699-34 Sequence 34, Appli  
42 182 12.1 810 3 US-09-273-565-34 Sequence 34, Appli  
43 182 12.1 810 4 US-09-565-538-34 Sequence 34, Appli  
44 182 12.1 810 4 US-09-661-468-34 Sequence 34, Appli  
45 182 12.1 810 4 US-09-976-165-34 Sequence 34, Appli

#### ALIGNMENTS

RESULT 1  
US-09-724-864-52  
; Sequence 52, Application US/09724864  
; Patent No. 6380362  
; GENERAL INFORMATION:  
; APPLICANT: Watson, James D.  
; APPLICANT: Murison, James G.  
; TITLE OF INVENTION: Polynucleotides, polypeptides expressed  
; FILE OF INVENTION: by the polynucleotides and methods for their use.  
; FILE REFERENCE: 11000.1050U1  
; CURRENT APPLICATION NUMBER: US/09/724,864  
; CURRENT FILING DATE: 2000-11-28  
; PRIOR APPLICATION NUMBER: U.S. No. 6380362 60/171,678  
; PRIOR FILING DATE: 1999-12-23  
; NUMBER OF SEQ ID NOS: 72  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 52  
; LENGTH: 278  
; TYPE: PRT  
; ORGANISM: Mouse  
US-09-724-864-52

Query Match 78.4%; Score 1180; DB 4; Length 278;  
Best Local Similarity 77.7%; Pred. No. 1.4e-86;  
Matches 213; Conservative 19; Mismatches 40; Indels 2; Gaps 2;  
QY 1 MRGSEVLLMWLLVAVGG-TEHAYRPGRRVCAYRAHGDVPVSEFVQVYQPFLLTCDGH 59  
Db 4 MWGSGELLVAVFLVLAADGTTTEHVYRPRRRVCTTGISGGSISSEFVQVYQPFLLTCDGH 63  
QY 60 RACSTYRTYRTAYRRSPQLAPARPRYACCPGWKRTSLPGACGAAICQPPCRNGGSCVQ 119  
Db 64 RACSTYRTYRTAYRRSPGVTPARPRYACCPGWKRTSLPGACGAAICQPPCRNGGSCIR 123  
QY 120 PRCRCPCAGWRGDTQSDVDECSARGCGPCRCINTAGSYWCQWEGHSLGADGTLCTPK 179  
Db 124 PGHCRCPGVWQGDTCQTDVDECSCTGEASCPCRCVNTVSGYWCQWEGSPSADGTRCLSK 183  
QY 180 GGPFRVAPNPT-GVDSAMKEEVORLQSRVDLLEKQLVLAPLHLSAOLEHGLPDPGS 238  
Db 184 EGSPFVAPNPTAGVDSAMKEEVYRQARVDVLEQLVLAPLHLSASRSTEHGLQDPGS 243  
QY 239 LLVHFSQQLGRIDSLSEQISFLEQLGSCSKKD 272  
Db 244 LLAHSFQQLDRIDSLSEQVSLFEEHLGSCSKKD 277

RESULT 2  
US-09-249-697A-6  
; Sequence 6, Application US/09249697A

KW colon tumour; breast tumour; prostate tumour; rectal tumour;  
KW liver tumour; bone disorder; cartilage disorder; sports injury;  
KW arthritis; wound.  
XX  
OS Homo sapiens.  
XX  
PN US2003045687-A1.  
XX  
PD 06-MAR-2003.  
XX  
PF 12-AUG-2002; 2002US-00218631.  
XX  
PR 01-JUN-2001; 2001WO-US017800.  
PR 29-JUN-2001; 2001WO-US021066.  
PR 09-APR-2002; 2002US-00119480.  
XX  
PA (GETH ) GENENTECH INC.  
XX  
PI Baker KP, Desnoyers L, Gerritsen ME, Goddard A, Godowski PJ;  
PI Grimaldi JC, Gurney AL, Smith V, Stephan JF, Watanabe CK, Wood WI;  
XX  
DR WPI; 2003-512315/48.  
DR N-PSDB; ACD68681.  
XX  
XX New genes, and its encoded secreted and transmembrane polypeptides,  
PT useful for stimulating Tumor Necrosis Factor alpha, or chondrocyte or  
PT pericyte proliferation, especially for treating lung tumors, arthritis or  
PT wounds in a mammal.  
XX  
PS Claim 11; Fig 178; 314pp; English.

XX The invention describes an isolated nucleic acid molecule comprising a  
CC sequence with at least 80% identity to: (a) a nucleotide encoding any of  
CC 122 PRO (secreted and transmembrane) polypeptides whose sequences are  
CC fully defined in the specification; or (b) any of 122 nucleotide  
CC sequences having e.g. 4834, 2504 or 1759 bp fully defined in the  
CC specification; or the full length coding sequence of any these 122  
CC nucleotide sequences. The PRO polypeptides or polynucleotides are useful  
CC as pharmaceuticals, diagnostics, biosensors or bioreactors. These are  
CC particularly useful for detecting tumors (e.g. lung tumour, colon  
CC tumour, breast tumour, prostate tumour, rectal tumour, or liver tumour)  
CC in a mammal, for stimulating the release of TNF-alpha from human blood,  
CC for stimulating the proliferation or differentiation of chondrocyte  
CC cells, for stimulating proliferation of pericyte cells, or for modulating  
CC normal human dermal fibroblast proliferation. The PRO nucleic acid or  
CC polypeptide is also useful for treating tumors or various bone and/or  
CC cartilage disorders (e.g. sports injuries or arthritis), or wounds. The  
CC PRO polypeptides are useful in drug screening, particularly as targets  
CC for therapeutic intervention in these diseases, and in the diagnostic  
CC determination of the presence of these diseases. The PRO polypeptides are  
CC also useful as molecular weight markers, or for chromosome  
CC identification. The PRO genes are useful as hybridisation probes, or for  
CC screening libraries of human cDNA, genomic DNA or mRNA. The PRO genes may  
CC also be used in gene therapy, particularly for replacing a defective  
CC gene. This is the amino acid sequence of a novel human secreted and  
CC transmembrane PRO polypeptide

XX Sequence 273 AA;

Query Match 100.0%; Score 1505; DB 6; Length 273;  
Best Local Similarity 100.0%; Pred. No. 1.3e-93;  
Matches 273; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MRGSQVLLMLLVAVGTEHAYRGRVCAVRAHGDVSVSFVQRVYQPELTTCDGHR 60  
DB 1 MRGSQVLLMLLVAVGTEHAYRGRVCAVRAHGDVSVSFVQRVYQPELTTCDGHR 60  
QY 61 ACSTYTYRTAYRRSFGAPAPRYACCPGWKRTSGLPGACGAACQPPCRNGSCVQP 120  
DB 61 ACSTYTYRTAYRRSFGAPAPRYACCPGWKRTSGLPGACGAACQPPCRNGSCVQP 120  
QY 121 GRRCPCPAGWRGDTCCSDVDECSARRGGCPQRCINTAGSYWCQWEGHSLSDGTLCVPKG 180

Db 121 GRRCPCPAGWRGDTCCSDVDECSARRGGCPQRCINTAGSYWCQWEGHSLSDGTLCVPKG 180  
QY 181 GPRVAPNPTGVDSAMKEEVQRLQSRVDLLEKQLQVLAPLHSLASQALEHGLPDPGSL 240  
Db 181 GPRVAPNPTGVDSAMKEEVQRLQSRVDLLEKQLQVLAPLHSLASQALEHGLPDPGSL 240  
QY 241 VHSFQQLGRIDSLSLSEQISFLEBQLGSCSCCKDS 273  
Db 241 VHSFQQLGRIDSLSLSEQISFLEBQLGSCSCCKDS 273

Search completed: April 22, 2004, 12:45:20  
Job time : 63 secs



PR 13-MAY-1998; 98US-00853339P.  
PR 15-MAY-1998; 98US-0085573P.  
PR 15-MAY-1998; 98US-0085579P.  
PR 15-MAY-1998; 98US-0085580P.  
PR 15-MAY-1998; 98US-0085582P.  
PR 15-MAY-1998; 98US-0085689P.  
PR 15-MAY-1998; 98US-0085697P.  
PR 15-MAY-1998; 98US-0085700P.  
PR 15-MAY-1998; 98US-0085704P.  
PR 18-MAY-1998; 98US-0086023P.  
PR 22-MAY-1998; 98US-0086392P.  
PR 22-MAY-1998; 98US-0086444P.  
PR 22-MAY-1998; 98US-0086430P.  
PR 22-MAY-1998; 98US-0086486P.  
PR 28-MAY-1998; 98US-0087098P.  
PR 28-MAY-1998; 98US-0087106P.  
PR 28-MAY-1998; 98US-0087280P.  
PR 26-JUN-1998; 98US-00105413.  
PR 26-JUN-1998; 98US-0090863P.  
PR 26-JUN-1998; 98US-0091010P.  
PR 01-JUL-1998; 98US-0091359P.  
PR 30-JUL-1998; 98US-0094651P.  
PR 11-SEP-1998; 98US-0100038P.  
PR 07-OCT-1998; 98US-00168978.  
PR 07-OCT-1998; 98WO-US021141.  
PR 02-NOV-1998; 98US-00184216.  
PR 06-NOV-1998; 98US-00187368.  
PR 20-NOV-1998; 98US-0109304P.  
PR 20-NOV-1998; 98WO-US024855.  
PR 07-DEC-1998; 98US-00202054.  
PR 22-DEC-1998; 98US-00218517.  
PR 23-DEC-1998; 98US-0113296P.  
PR 05-JAN-1999; 98WO-US000106.  
PR 05-MAR-1999; 99US-00254465.  
PR 08-MAR-1999; 99WO-US005028.  
PR 10-MAR-1999; 99US-00265686.  
PR 12-MAR-1999; 99WO-US005190.  
PR 12-MAR-1999; 99US-00267213.  
PR 12-MAR-1999; 99US-0123957P.  
PR 29-MAR-1999; 99US-0126773P.  
PR 12-APR-1999; 99US-00284291.  
PR 21-APR-1999; 99US-0130232P.  
PR 26-APR-1999; 99US-0131032P.  
PR 28-APR-1999; 99US-0131145P.  
PR 14-MAY-1999; 99US-00311832.  
PR 14-MAY-1999; 99US-0134287P.  
PR 02-JUN-1999; 99WO-US010733.  
PR 16-JUN-1999; 99US-0139557P.  
PR 07-JUL-1999; 99US-0141037P.  
PR 26-JUL-1999; 99US-0145698P.  
PR 28-JUL-1999; 99US-0146222P.  
PR 25-AUG-1999; 99US-00380137.  
PR 25-AUG-1999; 99US-00380138.  
PR 29-OCT-1999; 99US-0162506P.  
PR 30-NOV-1999; 99WO-US028213.  
PR 02-DEC-1999; 99WO-US028551.  
PR 02-DEC-1999; 99WO-US028565.  
PR 16-DEC-1999; 99WO-US030095.  
PR 30-DEC-1999; 99WO-US031243.  
PR 30-DEC-1999; 99WO-US031274.  
PR 05-JAN-2000; 2000WO-US000219.  
PR 06-JAN-2000; 2000WO-US000277.  
PR 06-JAN-2000; 2000WO-US000376.  
PR 11-FEB-2000; 2000WO-US003565.  
PR 18-FEB-2000; 2000WO-US004341.  
PR 24-FEB-2000; 2000WO-US005004.  
PR 02-MAR-2000; 2000WO-US005841.  
PR 10-MAR-2000; 2000WO-US006319.  
PR 21-MAR-2000; 2000WO-US007532.

PR 30-MAR-2000; 2000WO-US008439.  
PR 17-MAY-2000; 2000WO-US013705.  
PR 22-MAY-2000; 2000WO-US014042.  
PR 30-MAY-2000; 2000WO-US014941.  
PR 02-JUN-2000; 2000WO-US015264.  
PR 28-JUL-2000; 2000WO-US020710.  
PR 04-AUG-2000; 2000WO-US023328.  
PR 08-NOV-2000; 2000US-00709238.  
PR 27-NOV-2000; 2000US-00723749.  
PR 01-DEC-2000; 2000WO-US032678.  
PR 20-DEC-2000; 2000US-00747259.  
PR 28-FEB-2001; 2000WO-US034956.  
PR 22-MAR-2001; 2001WO-US006520.  
PR 22-MAR-2001; 2001US-00816744.  
PR 22-MAR-2001; 2001US-00816920.  
PR 22-MAR-2001; 2001WO-US009552.  
PR 10-MAY-2001; 2001US-00854208.  
PR 10-MAY-2001; 2001US-00854280.  
PR 25-MAY-2001; 2001WO-US017092.  
PR 01-JUN-2001; 2001US-00872035.  
PR 01-JUN-2001; 2001WO-US017800.  
PR 05-JUN-2001; 2001US-00874503.  
PR 14-JUN-2001; 2001US-00882636.  
PR 19-JUN-2001; 2001US-00886342.  
PR 20-JUN-2001; 2001WO-US019692.  
PR 29-JUN-2001; 2001WO-US021066.  
PR 09-JUL-2001; 2001WO-US021735.  
PR 30-JUL-2001; 2001US-00918585.  
XX  
XX (GETH ) GENENTECH INC.  
PA  
XX  
PI Ashkenazi AJ, Baker KP, Botstein D, Desnoyers L, Eaton DJ,  
PI Ferrara N, Filvaroff E, Fong S, Gao W, Gerber H, Gerritsen ME;

Query Match 100.0%; Score 1505; DB 6; Length 273;  
Best Local Similarity 100.0%; Pred. No. 1.3e-93;

Matches 273; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRGSQEVLLMVLVLAAGTGHAYRPGRRVCAVRAHGDVPSFVQRYVQPLTTCDGHR 60  
DB 1 MRGSQEVLLMVLVLAAGTGHAYRPGRRVCAVRAHGDVPSFVQRYVQPLTTCDGHR 60  
QY 61 ACSTYRTIYRTAYRSPGLAPRPRYACPGWKRTSGLPACGAACOPPCNGSCVQP 120  
DB 61 ACSTYRTIYRTAYRSPGLAPRPRYACPGWKRTSGLPACGAACOPPCNGSCVQP 120  
QY 121 GRCPAGWRGDTCSQSDVDECSARRGGPQRCINTAGSYWCQCEGHSLSDGTILCVPKG 180  
DB 121 GRCPAGWRGDTCSQSDVDECSARRGGPQRCINTAGSYWCQCEGHSLSDGTILCVPKG 180  
QY 181 GPRVAPNPTGVDSAMKEEVQRLQSRVDLLEBKQLVLAPLHSLASQALEHGLPPGSL 240  
DB 181 GPRVAPNPTGVDSAMKEEVQRLQSRVDLLEBKQLVLAPLHSLASQALEHGLPPGSL 240  
QY 241 VHSFOQLGRIDSLSEQISFLEELGSCSCCKDS 273  
DB 241 VHSFOQLGRIDSLSEQISFLEELGSCSCCKDS 273

## RESULT 15

ABO33793

ID ABO33793 standard; protein; 273 AA.

XX

AC ABO33793;

XX

DT 17-SEP-2003 (first entry)

XX

DE Novel human secreted and transmembrane protein PRO1330.

XX

KW Human; secreted and transmembrane protein; PRO; cytotostatic;

KW antiarthritic; osteopathic; gene therapy; TNF-Agonist-Alpha;

KW chondrocyte stimulator; pericyte stimulator; fibroblast modulator;

KW pharmaceutical; diagnostic; biosensor; bioreactor; tumour; lung tumour;

```
PR 01-DEC-2000; 2000WO-US032678.
PR 20-DEC-2000; 2000WO-US032678.
PR 20-DEC-2000; 2000WO-US032678.
PR 28-FEB-2001; 2001WO-US006520.
PR 22-MAR-2001; 2001US-00816744.
PR 22-MAR-2001; 2001US-00816920.
PR 22-MAR-2001; 2001WO-US009552.
PR 10-MAY-2001; 2001US-00854208.
PR 25-MAY-2001; 2001US-00854280.
PR 01-JUN-2001; 2001WO-US017092.
PR 01-JUN-2001; 2001WO-US017092.
PR 01-JUN-2001; 2001WO-US017092.
PR 05-JUN-2001; 2001US-00874503.
PR 14-JUN-2001; 2001US-00882636.
PR 19-JUN-2001; 2001US-00886342.
PR 20-JUN-2001; 2001WO-US019692.
PR 29-JUN-2001; 2001WO-US021066.
PR 09-JUL-2001; 2001WO-US021735.
PR 30-JUL-2001; 2001US-00918585.
XX
PA (GETH ) GENENTECH INC.
XX
PI Ashkenazi AJ, Baker KP, Botstein D, Desnoyers L, Eaton DL;
PI Ferrara N, Filvaroff E, Fong S, Gerber H, Gerritsen ME;
Query Match 100.0%; Score 1505; DB 6; Length 273;
Best Local Similarity 100.0%; Pred. No. 1.3e-93;
Matches 273; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MRGQEVLLMWLLVANGTTHAYRPGRRVCAVRAHGDVPSFVQVYQFFLTCDGHR 60
Db 1 MRGQEVLLMWLLVANGTTHAYRPGRRVCAVRAHGDVPSFVQVYQFFLTCDGHR 60
QY 61 ACSTYRTITRTYRSPGLAPRPRYACCPGWKTSGLPGACGAAIQQPCRRNGGSCVQP 120
Db 61 ACSTYRTITRTYRSPGLAPRPRYACCPGWKTSGLPGACGAAIQQPCRRNGGSCVQP 120
QY 121 GRCPAGWRGDTQSDVDECSARRGCGPQRCINTAGSYWCQWEGHSLADGTLCPKPG 180
Db 121 GRCPAGWRGDTQSDVDECSARRGCGPQRCINTAGSYWCQWEGHSLADGTLCPKPG 180
QY 181 GPPVAPNPTGVDAMKEEVORLSRVLDLLEKQLVLAHLASQALEHGLPDPGSL 240
Db 181 GPPVAPNPTGVDAMKEEVORLSRVLDLLEKQLVLAHLASQALEHGLPDPGSL 240
QY 241 VHSFOQLGRIDSLSEQISFLBEQLGSCSKKDS 273
Db 241 VHSFOQLGRIDSLSEQISFLBEQLGSCSKKDS 273
RESULT 14
ABO25272
ID ABO25272 standard; protein; 273 AA.
XX
AC ABO25272;
XX
DT 09-SEP-2003 (first entry)
XX
DE Novel human secreted and transmembrane protein PRO1330.
XX
KW Human; secreted and transmembrane protein; PRO; virucide; gene therapy;
KW cell death; growth induction cascade; blood coagulation cascade;
KW viral infection.
XX
OS Homo sapiens.
XX
PN US2003050239-A1.
XX
PD 13-MAR-2003.
XX
PF 15-OCT-2001; 2001US-00978191.
XX
PR 17-OCT-1997; 97US-0062250P.
PR 03-NOV-1997; 97US-0064249P.
PR 13-NOV-1997; 97US-0065311P.
PR 21-NOV-1997; 97US-0066364P.
PR 10-MAR-1998; 98US-0077450P.
PR 11-MAR-1998; 98US-0077632P.
PR 11-MAR-1998; 98US-0077641P.
PR 11-MAR-1998; 98US-0077649P.
PR 12-MAR-1998; 98US-0077791P.
PR 13-MAR-1998; 98US-0078004P.
PR 17-MAR-1998; 98US-00040220.
PR 20-MAR-1998; 98US-0078866P.
PR 20-MAR-1998; 98US-0078910P.
PR 20-MAR-1998; 98US-0078936P.
PR 20-MAR-1998; 98US-0078939P.
PR 25-MAR-1998; 98US-0079294P.
PR 26-MAR-1998; 98US-0079658P.
PR 27-MAR-1998; 98US-0079663P.
PR 27-MAR-1998; 98US-0079664P.
PR 27-MAR-1998; 98US-0079689P.
PR 27-MAR-1998; 98US-0079728P.
PR 27-MAR-1998; 98US-0079786P.
PR 30-MAR-1998; 98US-0079920P.
PR 30-MAR-1998; 98US-0079923P.
PR 31-MAR-1998; 98US-0080105P.
PR 31-MAR-1998; 98US-0080107P.
PR 31-MAR-1998; 98US-0080165P.
PR 31-MAR-1998; 98US-0080194P.
PR 01-APR-1998; 98US-0080327P.
PR 01-APR-1998; 98US-0080328P.
PR 01-APR-1998; 98US-0080333P.
PR 01-APR-1998; 98US-0080334P.
PR 08-APR-1998; 98US-0081049P.
PR 08-APR-1998; 98US-0081070P.
PR 08-APR-1998; 98US-0081071P.
PR 09-APR-1998; 98US-0081195P.
PR 09-APR-1998; 98US-0081203P.
PR 09-APR-1998; 98US-0081229P.
PR 15-APR-1998; 98US-0081817P.
PR 15-APR-1998; 98US-0081819P.
PR 15-APR-1998; 98US-0081838P.
PR 15-APR-1998; 98US-0081952P.
PR 15-APR-1998; 98US-0081955P.
PR 21-APR-1998; 98US-0082569P.
PR 21-APR-1998; 98US-0082569P.
PR 22-APR-1998; 98US-0082700P.
PR 22-APR-1998; 98US-0082704P.
PR 22-APR-1998; 98US-0082797P.
PR 22-APR-1998; 98US-0082804P.
PR 23-APR-1998; 98US-0082796P.
PR 27-APR-1998; 98US-0083336P.
PR 28-APR-1998; 98US-0083322P.
PR 29-APR-1998; 98US-0083392P.
PR 29-APR-1998; 98US-0083495P.
PR 29-APR-1998; 98US-0083496P.
PR 29-APR-1998; 98US-0083499P.
PR 29-APR-1998; 98US-0083500P.
PR 29-APR-1998; 98US-0083545P.
PR 29-APR-1998; 98US-0083554P.
PR 29-APR-1998; 98US-0083558P.
PR 29-APR-1998; 98US-0083559P.
PR 30-APR-1998; 98US-0083742P.
PR 05-MAY-1998; 98US-0084366P.
PR 06-MAY-1998; 98US-0084414P.
PR 06-MAY-1998; 98US-0084441P.
PR 07-MAY-1998; 98US-0084598P.
PR 07-MAY-1998; 98US-0084600P.
PR 07-MAY-1998; 98US-0084627P.
PR 07-MAY-1998; 98US-0084639P.
PR 07-MAY-1998; 98US-0084639P.
PR 07-MAY-1998; 98US-0084640P.
PR 07-MAY-1998; 98US-0084643P.
PR 13-MAY-1998; 98US-0085323P.
PR 13-MAY-1998; 98US-0085338P.
```

PR 17-MAR-1998; 98US-00040220.  
PR 20-MAR-1998; 98US-0078886P.  
PR 20-MAR-1998; 98US-0078910P.  
PR 20-MAR-1998; 98US-0078936P.  
PR 20-MAR-1998; 98US-0078939P.  
PR 25-MAR-1998; 98US-0079234P.  
PR 26-MAR-1998; 98US-0079656P.  
PR 27-MAR-1998; 98US-0079663P.  
PR 27-MAR-1998; 98US-0079664P.  
PR 27-MAR-1998; 98US-0079689P.  
PR 27-MAR-1998; 98US-0079728P.  
PR 27-MAR-1998; 98US-0079786P.  
PR 30-MAR-1998; 98US-0079920P.  
PR 30-MAR-1998; 98US-0079923P.  
PR 31-MAR-1998; 98US-0080105P.  
PR 31-MAR-1998; 98US-0080107P.  
PR 31-MAR-1998; 98US-0080165P.  
PR 31-MAR-1998; 98US-0080194P.  
PR 01-APR-1998; 98US-0080327P.  
PR 01-APR-1998; 98US-0080328P.  
PR 01-APR-1998; 98US-0080333P.  
PR 01-APR-1998; 98US-0080334P.  
PR 08-APR-1998; 98US-0081049P.  
PR 08-APR-1998; 98US-0081070P.  
PR 08-APR-1998; 98US-0081071P.  
PR 09-APR-1998; 98US-0081195P.  
PR 09-APR-1998; 98US-0081196P.  
PR 09-APR-1998; 98US-0081203P.  
PR 09-APR-1998; 98US-0081223P.  
PR 15-APR-1998; 98US-0081817P.  
PR 15-APR-1998; 98US-0081819P.  
PR 15-APR-1998; 98US-0081838P.  
PR 15-APR-1998; 98US-0081952P.  
PR 15-APR-1998; 98US-0081955P.  
PR 21-APR-1998; 98US-0082568P.  
PR 21-APR-1998; 98US-0082569P.  
PR 22-APR-1998; 98US-0082700P.  
PR 22-APR-1998; 98US-0082704P.  
PR 22-APR-1998; 98US-0082797P.  
PR 22-APR-1998; 98US-0082804P.  
PR 23-APR-1998; 98US-0082796P.  
PR 27-APR-1998; 98US-0083336P.  
PR 28-APR-1998; 98US-0083322P.  
PR 29-APR-1998; 98US-0083392P.  
PR 29-APR-1998; 98US-0083495P.  
PR 29-APR-1998; 98US-0083496P.  
PR 29-APR-1998; 98US-0083499P.  
PR 29-APR-1998; 98US-0083500P.  
PR 29-APR-1998; 98US-0083545P.  
PR 29-APR-1998; 98US-0083554P.  
PR 29-APR-1998; 98US-0083558P.  
PR 29-APR-1998; 98US-0083559P.  
PR 30-APR-1998; 98US-0083742P.  
PR 05-MAY-1998; 98US-0084366P.  
PR 06-MAY-1998; 98US-0084414P.  
PR 06-MAY-1998; 98US-0084443P.  
PR 07-MAY-1998; 98US-0084598P.  
PR 07-MAY-1998; 98US-0084600P.  
PR 07-MAY-1998; 98US-0084627P.  
PR 07-MAY-1998; 98US-0084637P.  
PR 07-MAY-1998; 98US-0084639P.  
PR 07-MAY-1998; 98US-0084640P.  
PR 07-MAY-1998; 98US-0084643P.  
PR 13-MAY-1998; 98US-0085323P.  
PR 13-MAY-1998; 98US-0085338P.  
PR 13-MAY-1998; 98US-0085339P.  
PR 15-MAY-1998; 98US-0085573P.  
PR 15-MAY-1998; 98US-0085579P.  
PR 15-MAY-1998; 98US-0085580P.  
PR 15-MAY-1998; 98US-0085582P.  
PR 15-MAY-1998; 98US-0085689P.  
PR 15-MAY-1998; 98US-0085697P.  
PR 15-MAY-1998; 98US-0085700P.  
PR 15-MAY-1998; 98US-0085704P.  
PR 18-MAY-1998; 98US-0086023P.  
PR 22-MAY-1998; 98US-0086392P.  
PR 22-MAY-1998; 98US-0086414P.  
PR 22-MAY-1998; 98US-0086430P.  
PR 22-MAY-1998; 98US-0086466P.  
PR 22-MAY-1998; 98US-0087098P.  
PR 28-MAY-1998; 98US-0087108P.  
PR 28-MAY-1998; 98US-0087208P.  
PR 26-JUN-1998; 98US-00105413.  
PR 26-JUN-1998; 98US-0090863P.  
PR 26-JUN-1998; 98US-0091010P.  
PR 01-JUL-1998; 98US-0091359P.  
PR 10-JUL-1998; 98US-0094651P.  
PR 11-SEP-1998; 98US-0100038P.  
PR 07-OCT-1998; 98US-00168978.  
PR 07-OCT-1998; 98US-0021141.  
PR 02-NOV-1998; 98US-00184216.  
PR 06-NOV-1998; 98US-00187368.  
PR 20-NOV-1998; 98US-0109304P.  
PR 20-NOV-1998; 98US-0109304P.  
PR 07-DEC-1998; 98US-00262054.  
PR 22-DEC-1998; 98US-00218517.  
PR 22-DEC-1998; 98US-0113296P.  
PR 23-DEC-1998; 98US-0113621P.  
PR 05-JAN-1999; 98US-0000106.  
PR 05-JAN-1999; 98US-00254465.  
PR 08-MAR-1999; 98US-0005028.  
PR 08-MAR-1999; 98US-00256686.  
PR 10-MAR-1999; 98US-00256686.  
PR 10-MAR-1999; 98US-0005190.  
PR 12-MAR-1999; 98US-00267213.  
PR 12-MAR-1999; 98US-0123957P.  
PR 29-MAR-1999; 98US-0126773P.  
PR 12-APR-1999; 98US-00284291.  
PR 21-APR-1999; 98US-0130232P.  
PR 26-APR-1999; 98US-0131022P.  
PR 28-APR-1999; 98US-0131445P.  
PR 14-MAY-1999; 98US-00311832.  
PR 14-MAY-1999; 98US-0134287P.  
PR 14-MAY-1999; 98US-00310733.  
PR 02-JUN-1999; 98US-0012252.  
PR 16-JUN-1999; 98US-0139557P.  
PR 23-JUN-1999; 98US-0141037P.  
PR 07-JUL-1999; 98US-0146808P.  
PR 26-JUL-1999; 98US-0145698P.  
PR 28-JUL-1999; 98US-0146222P.  
PR 25-AUG-1999; 98US-00380137.  
PR 25-AUG-1999; 98US-00380138.  
PR 25-AUG-1999; 98US-00380142.  
PR 29-OCT-1999; 98US-0163506P.  
PR 30-NOV-1999; 98US-0028313.  
PR 02-DEC-1999; 98US-0028551.  
PR 02-DEC-1999; 98US-0028565.  
PR 16-DEC-1999; 98US-0030095.  
PR 30-DEC-1999; 98US-0031243.  
PR 30-DEC-1999; 98US-0031274.  
PR 05-JAN-2000; 2000US-0000219.  
PR 06-JAN-2000; 2000US-0000277.  
PR 06-JAN-2000; 2000US-0000376.  
PR 11-FEB-2000; 2000US-0003565.  
PR 18-FEB-2000; 2000US-0004341.  
PR 24-FEB-2000; 2000US-0005004.  
PR 24-FEB-2000; 2000US-0005004.  
PR 10-MAR-2000; 2000US-0005841.  
PR 20-MAR-2000; 2000US-0006319.  
PR 21-MAR-2000; 2000US-0007532.  
PR 30-MAR-2000; 2000US-0008439.  
PR 17-MAY-2000; 2000US-0013705.  
PR 22-MAY-2000; 2000US-0014042.  
PR 30-MAY-2000; 2000US-0014941.  
PR 02-JUN-2000; 2000US-0015264.  
PR 28-JUL-2000; 2000US-0020710.  
PR 24-AUG-2000; 2000US-0023328.  
PR 08-NOV-2000; 2000US-00709238.  
PR 27-NOV-2000; 2000US-00723749.

CC treating proliferative disorders, inflammatory disorders. TANGO-175, or  
CC WDM-2 activity also include apoptotic disorders, rheumatoid arthritis,  
CC systemic lupus erythematosus, insulin-dependent diabetes mellitus, immune  
CC related disorders, e.g., immunodeficiency disorders, viral disorders,  
CC cell growth disorders, e.g., cancers and inflammatory disorders and  
CC apoptotic disorders. The nucleic acids of the invention are used in gene  
CC therapy. The present sequence is human T125 protein  
XX  
SQ Sequence 273 AA;

Query Match 100.0%; Score 1505; DB 5; Length 273;  
Best Local Similarity 100.0%; Pred. No. 1.3e-93;  
Matches 273; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRGSEVLLMLLVAVGCTEHAYRPGRRVCAVRAHGDVPSFVQRYQFLTTCDGHR 60  
DB 1 MRGSEVLLMLLVAVGCTEHAYRPGRRVCAVRAHGDVPSFVQRYQFLTTCDGHR 60  
QY 61 ACSTYRTTYRTAYRSPGLAPRPRYACCPGKRTSGLPACGAAICQPPCRNGSCVQP 120  
DB 61 ACSTYRTTYRTAYRSPGLAPRPRYACCPGKRTSGLPACGAAICQPPCRNGSCVQP 120  
QY 121 GRCRCFAGWRGDTCCSDVDECSARRGGCPQRCINTAGSYWCQWEGHSLSDGTLCPVKG 180  
DB 121 GRCRCFAGWRGDTCCSDVDECSARRGGCPQRCINTAGSYWCQWEGHSLSDGTLCPVKG 180  
QY 181 GPPRVAPNPTGYDSAMKEEVQRLQSRVLLLEKQLVLAPLHSLASQALEHGLDPGSL 240  
DB 181 GPPRVAPNPTGYDSAMKEEVQRLQSRVLLLEKQLVLAPLHSLASQALEHGLDPGSL 240  
QY 241 VHSFQOLGRIDSLSQISFLEELGSCCKKDS 273  
DB 241 VHSFQOLGRIDSLSQISFLEELGSCCKKDS 273

## RESULT 12

ABU0827  
ID ABU0827 standard; protein; 273 AA.  
XX  
AC ABU0827;  
XX  
DT 23-JUN-2003 (first entry)  
XX  
DE Human PRO polypeptide #89.  
XX  
XX Human; PRO polypeptide; secreted and transmembrane protein;  
KW anti-PRO antibody; diagnostic assay; gene expression; tumour; cytostatic.  
XX  
OS Homo sapiens.  
XX  
PN US2003036635-A1.  
XX  
PD 20-FEB-2003.  
XX  
PF 28-AUG-2002; 2002US-00230163.  
XX  
PR 25-JUL-2000; 2000US-0220638P.  
XX  
PR 01-JUN-2001; 2001WO-US017800.  
XX  
PR 29-JUN-2001; 2001WO-US021086.  
XX  
PR 09-APR-2002; 2002US-00119480.  
XX  
PA (GETH ) GENENTECH INC.  
XX  
PI Baker KP, Desnoyers L, Gerritsen ME, Goddard A, Godowski PJ;  
PI Grimaldi JC, Gurney AL, Smith V, Stephan JF, Watanabe CK, Wood WT;  
XX  
DR WPI; 2003-342045/32.  
DR N-PSDB; ACA66929.  
XX  
XX One hundred and twenty two nucleic acids encoding PRO polypeptides,  
PT useful for the manufacture of a medicament for diagnosing or treating  
PT tumor.  
PT  
XX

PS Claim 11; Fig 178; 314pp; English.  
XX  
CC The present invention relates to the isolation of novel human PRO  
CC polypeptides, and the polynucleotide sequences encoding them. The PRO  
CC polypeptides are secreted and transmembrane proteins. The PRO  
CC polypeptides and polynucleotides are useful for preparing a medicament  
CC useful in the diagnosis and treatment of tumours. Anti-PRO antibodies are  
CC useful in diagnostic assays for PRO, by detecting its expression in  
CC specific cells, tissues or serum, and for affinity purification of PRO  
CC from recombinant cell culture or natural sources. ABU0827-ABU0860  
CC represent the human PRO polypeptides of the invention. Note: The sequence  
CC data for this patent was obtained in electronic format directly from the  
CC USPTO web site at [seqdata.uspto.gov/psipadIDEntry.html](http://seqdata.uspto.gov/psipadIDEntry.html)  
XX

SQ Sequence 273 AA;

Query Match 100.0%; Score 1505; DB 6; Length 273;  
Best Local Similarity 100.0%; Pred. No. 1.3e-93;  
Matches 273; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRGSEVLLMLLVAVGCTEHAYRPGRRVCAVRAHGDVPSFVQRYQFLTTCDGHR 60  
DB 1 MRGSEVLLMLLVAVGCTEHAYRPGRRVCAVRAHGDVPSFVQRYQFLTTCDGHR 60  
QY 61 ACSTYRTTYRTAYRSPGLAPRPRYACCPGKRTSGLPACGAAICQPPCRNGSCVQP 120  
DB 61 ACSTYRTTYRTAYRSPGLAPRPRYACCPGKRTSGLPACGAAICQPPCRNGSCVQP 120  
QY 121 GRCRCFAGWRGDTCCSDVDECSARRGGCPQRCINTAGSYWCQWEGHSLSDGTLCPVKG 180  
DB 121 GRCRCFAGWRGDTCCSDVDECSARRGGCPQRCINTAGSYWCQWEGHSLSDGTLCPVKG 180  
QY 181 GPPRVAPNPTGYDSAMKEEVQRLQSRVLLLEKQLVLAPLHSLASQALEHGLDPGSL 240  
DB 181 GPPRVAPNPTGYDSAMKEEVQRLQSRVLLLEKQLVLAPLHSLASQALEHGLDPGSL 240  
QY 241 VHSFQOLGRIDSLSQISFLEELGSCCKKDS 273  
DB 241 VHSFQOLGRIDSLSQISFLEELGSCCKKDS 273

## RESULT 13

ABO25271  
ID ABO25271 standard; protein; 273 AA.  
XX  
AC ABO25271;  
XX  
DT 09-SEP-2003 (first entry)  
XX  
DE Novel human secreted and transmembrane protein PRO213-1.  
XX  
XX Human; secreted and transmembrane protein; PRO; virucide; gene therapy;  
KW cell death; growth induction cascade; blood coagulation cascade;  
KW viral infection.  
XX  
OS Homo sapiens.  
XX  
PN US2003050239-A1.  
XX  
PD 13-MAR-2003.  
XX  
PF 15-OCT-2001; 2001US-00978191.  
XX  
XX 17-OCT-1997; 97US-0062250P.  
XX 03-NOV-1997; 97US-0064249P.  
XX 13-NOV-1997; 97US-0065311P.  
XX 21-NOV-1997; 97US-0066364P.  
XX 10-MAR-1998; 98US-0077450P.  
XX 11-MAR-1998; 98US-0077632P.  
XX 11-MAR-1998; 98US-0077641P.  
XX 11-MAR-1998; 98US-0077649P.  
XX 12-MAR-1998; 98US-0077791P.  
XX 13-MAR-1998; 98US-0078004P.

PI Grimaldi JC, Gurney AL, Smith V, Stephan JP, Watanabe CK, Wood WI;  
 XX WPI; 2002-172001/22.  
 DR N-PSDB; ABK33624.  
 XX  
 PT One hundred and twenty two nucleic acids encoding PRO polypeptides,  
 PT useful for treating a PRO related disorder and for diagnosing tumors such  
 PT as lung cancer, colon cancer, breast tumor, prostate tumor, rectal tumor  
 PT or liver tumor.  
 XX  
 PS Claim 11; Fig 178; 359pp; English.

XX The invention relates to one hundred and twenty two nucleic acids  
 CC encoding PRO polypeptides. The sequences of the 122 PRO polynucleotides  
 CC encode human secreted proteins. The PRO nucleic acids, polypeptides,  
 CC agonists and antagonists are useful for treating a PRO related disorder.  
 CC The PRO polypeptides are useful for diagnosing tumours, especially lung  
 CC cancer, colon cancer, breast tumor, prostate tumor, rectal tumor or  
 CC liver tumor. The PRO polypeptides are useful for stimulating the  
 CC proliferation or differentiation of chondrocyte cells, for stimulating  
 CC the proliferation or release of tumour necrosis factor-alpha from human blood,  
 CC for stimulating or inhibiting the proliferation of normal human dermal  
 CC fibroblast cells. The PRO polypeptide may also be used as molecular  
 CC weight markers and for tissue typing. The PRO nucleic acids have  
 CC applications in molecular biology, including use as hybridisation probes,  
 CC and in chromosome and gene mapping. AAU83592-AAU83713 represent human PRO  
 CC protein sequences of the invention  
 XX  
 SQ Sequence 273 AA;

Query Match 100.0%; Score 1505; DB 5; Length 273;  
 Best Local Similarity 100.0%; Pred. No. 1.3e-93;  
 Matches 273; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRGSEVLLMWLLVAVGTEHAYRPGRRVCAVEAHGDPVSEFVQRYVQFFLTCDGHR 60  
 Db 1 MRGSEVLLMWLLVAVGTEHAYRPGRRVCAVEAHGDPVSEFVQRYVQFFLTCDGHR 60  
 QY 61 ACSTYRTIYTYRSPGLAPRPRYACCPGWKRTSLGPGACGAACQPPCRNGSGVQ 120  
 Db 61 ACSTYRTIYTYRSPGLAPRPRYACCPGWKRTSLGPGACGAACQPPCRNGSGVQ 120  
 QY 121 GRCEPAGWGDTCQSDVDCSARRGCPORCINTAGSYWCQWEGHSLADGTLCPVK 180  
 Db 121 GRCEPAGWGDTCQSDVDCSARRGCPORCINTAGSYWCQWEGHSLADGTLCPVK 180  
 QY 181 GPPRVAENPTGVDSEAMKEEVQRLQSRVDLLEKQLVLPLHSLASQALEHGLPDPGSL 240  
 Db 181 GPPRVAENPTGVDSEAMKEEVQRLQSRVDLLEKQLVLPLHSLASQALEHGLPDPGSL 240  
 QY 241 VHSFOQLGRDLSLEQISFLEQLGSCSKKDS 273  
 Db 241 VHSFOQLGRDLSLEQISFLEQLGSCSKKDS 273

RESULT 11

AAE21079

ID AAE21079 standard; protein; 273 AA.

XX

AC AAE21079;

XX

DT 01-JUL-2002 (first entry)

XX

DE Human T125 (TANGO-125) protein.

XX

KW Human; haematopoiesis; clotting; kidney failure; wound healing; cancer;  
 KW neoplasia; pancreatic disorder; pancreatitis; cerebrovascular disease;  
 KW heart disorder; ischaemic heart disease; neuroprotective; vulvectomy;  
 KW cardiovascular disorder; ischaemic heart disease; immunosuppressive;  
 KW glomerular disease; glomerulonephritis; uterine disorder; hyperplasia;  
 KW fetal spleen; prostate disorder; inflammatory disease; Crohn's disease;  
 KW Proliferative disorder; gynaecological; haemostatic; antibacterial;

KW systemic lupus erythematosus; immunodeficiency disorder; antiasthmatic;  
 KW cytostatic; nephrotropic; antidiabetic; cerebroprotective; tranquiliser;  
 KW hypotensive; tumour; injury; trauma; antiangiinal; vasotropic; antilucer;  
 KW apototic disorder; rheumatoid arthritis; cardiac; renal disorder;  
 KW hepatotropic; antipsoriatic; antiallergic; dermatological; virucide;  
 T125.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Peptide 1..22  
 FT Protein /label= Signal\_peptide  
 FT 23..273  
 FT Domain /note= "Human mature T125 protein"  
 FT 107..134  
 FT Domain /note= "Epidermal growth factor domain 1"  
 FT 141..176  
 FT Domain /note= "Epidermal growth factor domain 2"

US2002028508-A1.

07-MAR-2002.

21-FEB-2001; 2001US-00790254.

23-APR-1998; 98US-00065363.

23-APR-1998; 98US-00065661.

22-JUN-1998; 98US-00102705.

29-JUL-1998; 98US-00124538.

23-APR-1999; 98US-00298531.

22-JUN-1999; 98US-00337930.

29-JUL-1999; 98US-00363630.

(HOLT/) HOLTZMAN D A.

(GOOD/) GOODEARL A D J.

(MCCA/) MCCARTHY S A.

Holtzman DA, Goodearl ADJ, McCarthy SA;

WPI; 2002-303420/34.

N-PSDB; AAD33536.

Novel TANGO polypeptides and nucleic acid molecules useful as modulating  
 agents in regulating cellular processes and for diagnosing and treating  
 heart, liver, lung, kidney, inflammatory and cellular proliferative  
 disorders.

Claim 51; Fig 4; 138pp; English.

The invention relates to nucleic acids encoding a variety of proteins  
 human T139 (TANGO-139), T125 (TANGO-125), T110 (TANGO-110), murine T175  
 (TANGO-175), human T175 or murine WDM-2, having diagnostic, preventive,  
 therapeutic and other uses. Polypeptide of the invention has the ability  
 to inhibit a proteinase activity, to modulate cell-cell interactions,  
 haematopoiesis and the ability to modulate clotting. Polypeptide and  
 polynucleotide of the invention are useful for diagnosing and treating  
 disorder characterised by their aberrant expression or activity. The  
 antibodies are useful as modulating agents in regulating a variety of  
 cellular processes e.g. cell proliferation and/or cell differentiation.  
 TANGO-139 is useful for treating kidney defects such as kidney failure.  
 TANGO-125 is useful in wound healing and for treating cancer, TANGO-110  
 is useful for treating neoplasia, TANGO-177 or WDM-2 is useful for  
 treating cancer, are useful to treat pancreatic disorders, such as  
 pancreatitis, cerebrovascular disease, and tumours, and injury or trauma  
 to the brain. TANGO-125, 110, 175 molecules treat heart disorders, e.g.,  
 ischaemic heart disease, cardiovascular disorders, such as ischaemic  
 heart disease. TANGO-139, 125, 110 and 175 molecules are useful to treat  
 renal (kidney) disorders, such as glomerular disease (e.g., acute and  
 chronic glomerulonephritis), TANGO-175 is useful to treat uterine  
 disorders, hyperplasia of the endometrium. TANGO-110 is useful to treat  
 spleen, e.g., the fetal spleen, associated diseases and disorder. TANGO-  
 125 treats prostate disorders, such as inflammatory diseases, Crohn's  
 disease and tumours. TANGO-139, 125, 110, 175 or WDM-2 are useful for

RESULT 9  
 AAB24043  
 ID AAB24043 standard; protein; 273 AA.  
 AC  
 AC AAB24043;  
 XX  
 XX  
 DT 25-JAN-2001 (first entry)  
 XX  
 XX Human PRO1330 protein sequence SEQ ID NO.6.  
 DE  
 XX Human; tumour; diagnosis; neoplastic disease; identification; cancer;  
 KW tumorigenesis; detection; neoplastic cell growth; proliferation;  
 KW cytostatic; antiinflammatory; immunomodulatory; inflammatory disorder;  
 KW immunological disorder.  
 XX  
 XX Homo sapiens.  
 OS  
 XX WO200053754-A1.  
 PN  
 XX  
 PD 14-SEP-2000.  
 XX  
 XX 06-JAN-2000; 2000WO-US000277.  
 PF  
 XX 08-MAR-1999; 99WO-US005028.  
 PR 12-MAR-1999; 99US-0123957P.  
 PR 29-MAR-1999; 99US-0126773P.  
 PR 21-APR-1999; 99US-0130232P.  
 PR 28-APR-1999; 99US-0131445P.  
 PR 05-OCT-1999; 99WO-US023089.  
 PR 30-NOV-1999; 99WO-US028313.  
 PR 02-DEC-1999; 99WO-US028551.  
 PR 02-DEC-1999; 99WO-US028564.  
 PR 30-DEC-1999; 99WO-US031243.  
 PR 30-DEC-1999; 99WO-US031274.  
 XX  
 XX (GETH ) GENENTECH INC.  
 PA  
 XX Baker KP, Desauvage PJ, Goddard A, Gurney AL, Klein RD, Roy MA;  
 PI Wood WI;  
 XX  
 XX WPI; 2000-572269/53.  
 DR N-PSDB; AAC58227.  
 DR  
 XX  
 XX New isolated antibody for use in compositions and methods for the  
 PT diagnosis and treatment of neoplastic cell growth and proliferation in  
 PT mammals, including humans, and in monitoring tumor treatment.  
 XX  
 XX Claim 61; Fig 6; 195pp; English.  
 PS  
 XX The present invention describes an isolated antibody (Ab) that binds to  
 CC one of the human proteins (P) designated PRO213, PRO1330, PRO1449,  
 CC PRO237, PRO324, PRO351, PRO362, PRO615, PRO531, PRO538, PRO3664, PRO618,  
 CC PRO772, PRO703, PRO792 or PRO474. The Ab can be used in compositions and  
 CC methods for the diagnosis and treatment of neoplastic cell growth and  
 CC proliferation in mammals, including humans. Genes and polypeptides  
 CC encoded by them, that are amplified in the genome of a tumor cell, can  
 CC be identified and are useful targets for the treatment and prevention of  
 CC certain cancers and may be used to monitor tumor treatment. Compounds  
 CC that inhibit the expression or activity of the identified polypeptides  
 CC can be identified and used as antagonists. Benign or malignant tumors,  
 CC inflammatory disorders and immunological disorders can be treated.  
 CC AAC58123 to AAC58224 represent hybridisation probes and PCR primers used  
 CC in the isolation of the human PRO sequences. AAC58225 to AAC58241 and  
 CC AAB24041 to AAB24056 represent human PRO polynucleotide and protein  
 CC sequences given in the exemplification of the present invention  
 XX  
 XX Sequence 273 AA;  
 SQ  
 Query Match 100.0%; Score 1505; DB 3; Length 273;  
 Best Local Similarity 100.0%; Pred. No. 1.3e-93;  
 Matches 273; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 PI Baker KP, Desnoyers L, Gerritsen ME, Goddard A, Godowski PJ;  
 (GETH ) GENENTECH INC.  
 QY 1 MRGSEVLLMWLLVLAVGGTEHAYRGRVCAVRAHGDVPSBESFVQVYQPLTTCDGHR 60  
 DB 1 MRGSEVLLMWLLVLAVGGTEHAYRGRVCAVRAHGDVPSBESFVQVYQPLTTCDGHR 60  
 QY 61 ACSTYRTTYRTAYRRSPGLAPAPRPRVACCPGWKRTSGLPACGACGAAICQPPCRNGGSCVQP 120  
 DB 61 ACSTYRTTYRTAYRRSPGLAPAPRPRVACCPGWKRTSGLPACGACGAAICQPPCRNGGSCVQP 120  
 QY 121 GRCRCPAGWRGDTCCQSDVDECSARRGGCPQRCINTAGSYWCOCWEHSHLSADGTLCPVKG 180  
 DB 121 GRCRCPAGWRGDTCCQSDVDECSARRGGCPQRCINTAGSYWCOCWEHSHLSADGTLCPVKG 180  
 QY 181 GPRVPAPNPTGVDSAMKEEVQRIQSRVDLLEEKQLVLAPLHSLASQALEHGLPDPGSL 240  
 DB 181 GPRVPAPNPTGVDSAMKEEVQRIQSRVDLLEEKQLVLAPLHSLASQALEHGLPDPGSL 240  
 QY 241 VHSFQQLGRIDSLSEQISFLEOLGSCCKKDS 273  
 DB 241 VHSFQQLGRIDSLSEQISFLEOLGSCCKKDS 273  
 RESULT 10  
 AAU83680  
 ID AAU83680 standard; protein; 273 AA.  
 AC  
 AC AAU83680;  
 DT 08-MAY-2002 (first entry)  
 XX  
 XX Human PRO protein, Seq ID No 178.  
 DE  
 XX Human; secreted protein; PRO; tumour; lung cancer; colon cancer;  
 KW breast cancer; prostate tumour; rectal tumour; liver tumour;  
 KW pericyte cell proliferation; chondrocyte cell proliferation;  
 KW tumour necrosis factor-alpha.  
 XX  
 XX Homo sapiens.  
 OS  
 XX WO200208288-A2.  
 PN  
 XX 31-JAN-2002.  
 PD  
 XX 29-JUN-2001; 2001WO-US021066.  
 PF  
 XX 20-JUL-2000; 2000US-0219556P.  
 PR 25-JUL-2000; 2000US-0220585P.  
 PR 25-JUL-2000; 2000US-0220605P.  
 PR 25-JUL-2000; 2000US-0220607P.  
 PR 25-JUL-2000; 2000US-0220624P.  
 PR 25-JUL-2000; 2000US-0220638P.  
 PR 25-JUL-2000; 2000US-0220664P.  
 PR 25-JUL-2000; 2000US-0220666P.  
 PR 26-JUL-2000; 2000US-0220893P.  
 PR 28-JUL-2000; 2000WO-US020710.  
 PR 01-AUG-2000; 2000US-0222425P.  
 PR 22-AUG-2000; 2000US-0227133P.  
 PR 23-AUG-2000; 2000WO-US023522.  
 PR 24-AUG-2000; 2000WO-US033328.  
 PR 10-NOV-2000; 2000WO-US030873.  
 PR 28-NOV-2000; 2000US-0253646P.  
 PR 01-DEC-2000; 2000WO-US032678.  
 PR 20-DEC-2000; 2000US-00747259.  
 PR 28-FEB-2001; 2001WO-US004956.  
 PR 01-MAR-2001; 2001WO-US006520.  
 PR 22-MAR-2001; 2001US-00816744.  
 PR 10-MAY-2001; 2001US-00854208.  
 PR 10-MAY-2001; 2001US-00854280.  
 PR 25-MAY-2001; 2001WO-US017092.  
 XX  
 XX (GETH ) GENENTECH INC.  
 PA  
 XX Baker KP, Desnoyers L, Gerritsen ME, Goddard A, Godowski PJ;  
 PI

```
PR 14-MAY-1999; 99US-0134287P.
XX 02-DEC-1999; 99WO-US028565.
XX (GETH ) GENENTECH INC.
XX Baker KP, Ferrara N, Goddard A, Gurney AL, Hillan KJ;
XX Williams PM, Wood WI;
XX WPI; 2000-638138/61.
XX N-PSDB; AAA75703.
XX A composition useful for treatment and diagnosis of a cardiovascular,
XX endothelial or angiogenic disorder, especially cancer, comprises (an
XX agonist or antagonist of) a PRO320, PRO338, PRO1031, PRO296, PRO213,
XX PRO1330 or PRO1449 polypeptide.
XX Claim 67; Fig 12; 152pp; English.
XX
XX The present sequence represents PRO1330, a human notch 4 homologue. The
XX specification describes PRO320, PRO338, PRO1031, PRO296, PRO213, PRO1330
XX and PRO1449 polypeptides. The polypeptides promoter or inhibit
XX angiogenesis and cardiovascularisation in mammals. The polypeptides are
XX used for the treatment and diagnosis of a cardiovascular, endothelial or
XX angiogenic disorder, especially cancer. Disorders that can be diagnosed,
XX treated or prevented by the polypeptides of the invention include trauma
XX such as wounds, arteriosclerosis, and cardiac hypertrophy
XX
XX Sequence 273 AA;
XX
XX Query Match 100.0%; Score 1505; DB 3; Length 273;
XX Best Local Similarity 100.0%; Pred. No. 1.3e-93;
XX Matches 273; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 MRGQEVLLMVLAVGTEHAYRPGRRVCAVRAHGDVPSSEFVQRYVQPLTTCDGHR 60
XX DB 1 MRGQEVLLMVLAVGTEHAYRPGRRVCAVRAHGDVPSSEFVQRYVQPLTTCDGHR 60
XX
XX QY 61 ACSTYRTIYTRRSPGLAPRPYACCPGWKRTSGLPACGAAICQPCRNNGSCVQP 120
XX DB 61 ACSTYRTIYTRRSPGLAPRPYACCPGWKRTSGLPACGAAICQPCRNNGSCVQP 120
XX
XX QY 121 GRCPAGWRGDTQSDVDECSARRGCPQRCINTAGSYWCQWEGHSLADGTLCPVK 180
XX DB 121 GRCPAGWRGDTQSDVDECSARRGCPQRCINTAGSYWCQWEGHSLADGTLCPVK 180
XX
XX QY 181 GPPRVAPNPTGVDSAMKEEVQRLQSRVDLLEKQLVLAPLHSLASQALEHGLPDPGSL 240
XX DB 181 GPPRVAPNPTGVDSAMKEEVQRLQSRVDLLEKQLVLAPLHSLASQALEHGLPDPGSL 240
XX
XX QY 241 VHSFOQLGRIDSLSEQISFLEEQLGSCCKDS 273
XX DB 241 VHSFOQLGRIDSLSEQISFLEEQLGSCCKDS 273
XX
XX RESULT 8
XX AAB24042
XX ID AAB24042 standard; protein; 273 AA.
XX AC AAB24042;
XX XX
XX DT 25-JAN-2001 (first entry)
XX
XX DE Human PRO213 protein sequence SEQ ID NO:4.
XX
XX KW Human; tumour; diagnosis; neoplastic disease; identification; cancer;
XX KW tumorigenesis; detection; neoplastic cell growth; proliferation;
XX KW cytostatic; antiinflammatory; immunomodulatory; inflammatory disorder;
XX KW immunological disorder.
XX
XX OS Homo sapiens.
XX
XX FN WO2000053754-A1.
XX
```

```
PD 14-SEP-2000.
XX 06-JAN-2000; 2000WO-US000277.
XX
XX 08-MAR-1999; 99WO-US005028.
XX 12-MAR-1999; 99US-0123957P.
XX 29-MAR-1999; 99US-0126773P.
XX 21-APR-1999; 99US-0130232P.
XX 28-APR-1999; 99US-0131445P.
XX 05-OCT-1999; 99WO-US023089.
XX 30-NOV-1999; 99WO-US028313.
XX 02-DEC-1999; 99WO-US028551.
XX 02-DEC-1999; 99WO-US028564.
XX 30-DEC-1999; 99WO-US031243.
XX 30-DEC-1999; 99WO-US031274.
XX
XX (GETH ) GENENTECH INC.
XX
XX PI Baker KP, Desauvage FJ, Goddard A, Gurney AL, Klein RD, Roy MA;
XX Wood WI;
XX WPI; 2000-572369/53.
XX N-PSDB; AAC58226.
XX
XX New isolated antibody for use in compositions and methods for the
XX diagnosis and treatment of neoplastic cell growth and proliferation in
XX mammals, including humans, and in monitoring tumor treatment.
XX
XX Claim 61; Fig 4; 195pp; English.
XX
XX The present invention describes an isolated antibody (Ab) that binds to
XX one of the human proteins (P) designated PRO213, PRO1330, PRO1449,
XX PRO337, PRO324, PRO351, PRO362, PRO615, PRO531, PRO538, PRO3664, PRO618,
XX PRO772, PRO703, PRO792 or PRO474. The Ab can be used in compositions and
XX methods for the diagnosis and treatment of neoplastic cell growth and
XX proliferation in mammals, including humans. Genes and polypeptides
XX encoded by them, that are amplified in the genome of a tumour cell, can
XX be identified and are useful targets for the treatment and prevention of
XX certain cancers and may be used to monitor tumour treatment. Compounds
XX that inhibit the expression or activity of the identified polypeptides
XX can be identified and used as antagonists. Benign or malignant tumours,
XX inflammatory disorders and immunological disorders can be treated.
XX AAC58123 to AAC58224 represent hybridisation probes and PCR primers used
XX in the isolation of the human PRO sequences. AAC58225 to AAC58241 and
XX AAB24041 to AAB24056 represent human PRO polynucleotide and protein
XX sequences given in the exemplification of the present invention
XX
XX Sequence 273 AA;
XX
XX Query Match 100.0%; Score 1505; DB 3; Length 273;
XX Best Local Similarity 100.0%; Pred. No. 1.3e-93;
XX Matches 273; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 MRGQEVLLMVLAVGTEHAYRPGRRVCAVRAHGDVPSSEFVQRYVQPLTTCDGHR 60
XX DB 1 MRGQEVLLMVLAVGTEHAYRPGRRVCAVRAHGDVPSSEFVQRYVQPLTTCDGHR 60
XX
XX QY 61 ACSTYRTIYTRRSPGLAPRPYACCPGWKRTSGLPACGAAICQPCRNNGSCVQP 120
XX DB 61 ACSTYRTIYTRRSPGLAPRPYACCPGWKRTSGLPACGAAICQPCRNNGSCVQP 120
XX
XX QY 121 GRCPAGWRGDTQSDVDECSARRGCPQRCINTAGSYWCQWEGHSLADGTLCPVK 180
XX DB 121 GRCPAGWRGDTQSDVDECSARRGCPQRCINTAGSYWCQWEGHSLADGTLCPVK 180
XX
XX QY 181 GPPRVAPNPTGVDSAMKEEVQRLQSRVDLLEKQLVLAPLHSLASQALEHGLPDPGSL 240
XX DB 181 GPPRVAPNPTGVDSAMKEEVQRLQSRVDLLEKQLVLAPLHSLASQALEHGLPDPGSL 240
XX
XX QY 241 VHSFOQLGRIDSLSEQISFLEEQLGSCCKDS 273
XX DB 241 VHSFOQLGRIDSLSEQISFLEEQLGSCCKDS 273
XX
```





Query Match 100.0%; Score 1505; DB 3; Length 273;  
 Best Local Similarity 100.0%; Pred. No. 1.3e-93;  
 Matches 273; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRGQEVLLMVLAVGGTEHAYRPGRRVCAVRAHGDVPVSEFVQRYVQFLLTCDGHR 60  
 DB 1 MRGQEVLLMVLAVGGTEHAYRPGRRVCAVRAHGDVPVSEFVQRYVQFLLTCDGHR 60

QY 61 ACSTYRTIYTAVERSPGLAPAPRYACCPGWKRTSLGACGAAICQPPCRNGGSCVQP 120  
 DB 61 ACSTYRTIYTAVERSPGLAPAPRYACCPGWKRTSLGACGAAICQPPCRNGGSCVQP 120

QY 121 GRCCPAGMRGDTQSDVDECSARRGCGPCRCINTAGSYWCQWEGHSLADGTLCPVKG 180  
 DB 121 GRCCPAGMRGDTQSDVDECSARRGCGPCRCINTAGSYWCQWEGHSLADGTLCPVKG 180

QY 181 GPRVAPNPTGVDSAMKEEVQRLQSRVLDLLEKQLVLAPLHSLAQALEHGLPDPGSL 240  
 DB 181 GPRVAPNPTGVDSAMKEEVQRLQSRVLDLLEKQLVLAPLHSLAQALEHGLPDPGSL 240

QY 241 VHSFQQLGRIDSLSEQISFLEEQLGSCCKDS 273  
 DB 241 VHSFQQLGRIDSLSEQISFLEEQLGSCCKDS 273

RESULT 5  
 AAB44325  
 ID AAB44325 standard; protein; 273 AA.  
 AC AAB44325;  
 XX  
 DT 08-FEB-2001 (first entry)  
 DE Human PRO213-1 protein sequence SEQ ID NO:506.  
 KW Human; secreted protein; transmembrane protein; PRO; EST; cytostatic;  
 KW expressed sequence tag; detection; cancer.  
 XX Homo sapiens.  
 XX WO200053756-A2.  
 PD 14-SEP-2000.  
 XX  
 PF 18-FEB-2000; 2000WO-US0004341.  
 XX  
 PR 08-MAR-1999; 99WO-US0005028.  
 PR 12-MAR-1999; 99US-0123957P.  
 PR 29-MAR-1999; 99US-0126773P.  
 PR 21-APR-1999; 99US-0130232P.  
 PR 28-APR-1999; 99US-0131445P.  
 PR 14-MAY-1999; 99US-0134287P.  
 PR 23-JUN-1999; 99US-0141037P.  
 PR 26-JUL-1999; 99US-0145698P.  
 PR 29-OCT-1999; 99US-0162506P.  
 PR 30-NOV-1999; 99WO-US028313.  
 PR 02-DEC-1999; 99WO-US028551.  
 PR 16-DEC-1999; 99WO-US030095.  
 PR 30-DEC-1999; 99WO-US031243.  
 PR 30-DEC-1999; 99WO-US031274.  
 PR 05-JAN-2000; 2000WO-US000219.  
 PR 06-JAN-2000; 2000WO-US000277.  
 PR 06-JAN-2000; 2000WO-US000376.  
 XX  
 PA (SETH) GENENTECH INC.  
 XX  
 PI Ashkenazi AJ, Baker KP, Botstein D, Desnoyers L, Eaton DL;  
 PI Ferrara N, Filvaroff E, Fong S, Gao W, Gerber H, Gerritsen ME;  
 PI Goddard A, Godowski PJ, Grimaldi CJ, Gurney AL, Hillan KJ;  
 PI Kijavini IJ, Kuo SS, Napier MA, Pan J, Paoni NF, Roy WA, Shelton DL;  
 PI Stewart TA, Tumas D, Williams PM, Wood WI;  
 XX

DR WPI: 2000-611443/58.  
 DR N-PSDB; AAC78585.  
 XX Novel PRO polypeptides and polynucleotides used in detection methods, to  
 PT target bioactive molecules to specific cells, and to modulate cellular  
 PT activities.  
 XX Claim 12; Fig 213; 636pp; English.  
 XX AAC78458 to AAC78599 represent polynucleotide and EST (expressed sequence  
 CC tag) sequences which encode secreted or transmembrane PRO polypeptides.  
 CC The PRO polynucleotides and polypeptides have cytostatic activity. The  
 CC polynucleotides and polypeptides can be used for detecting the presence  
 CC of PRO polypeptides in samples, for linking bioactive molecules to cells  
 CC and for modulating biological activities of cells, using the polypeptides  
 CC for specific targeting. The polypeptide targeting can be used to kill the  
 CC target cells, e.g. for the treatment of cancers. The polypeptide pairs  
 CC provide specific targeting of bioactive molecules to cells. AAC78600 to  
 CC AAC78987 represent PCR primers and probes used in the isolation of the  
 CC PRO polynucleotide sequences  
 XX Sequence 273 AA;  
 SQ

Query Match 100.0%; Score 1505; DB 3; Length 273;  
 Best Local Similarity 100.0%; Pred. No. 1.3e-93;  
 Matches 273; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRGQEVLLMVLAVGGTEHAYRPGRRVCAVRAHGDVPVSEFVQRYVQFLLTCDGHR 60  
 DB 1 MRGQEVLLMVLAVGGTEHAYRPGRRVCAVRAHGDVPVSEFVQRYVQFLLTCDGHR 60

QY 61 ACSTYRTIYTAVERSPGLAPAPRYACCPGWKRTSLGACGAAICQPPCRNGGSCVQP 120  
 DB 61 ACSTYRTIYTAVERSPGLAPAPRYACCPGWKRTSLGACGAAICQPPCRNGGSCVQP 120

QY 121 GRCCPAGMRGDTQSDVDECSARRGCGPCRCINTAGSYWCQWEGHSLADGTLCPVKG 180  
 DB 121 GRCCPAGMRGDTQSDVDECSARRGCGPCRCINTAGSYWCQWEGHSLADGTLCPVKG 180

QY 181 GPRVAPNPTGVDSAMKEEVQRLQSRVLDLLEKQLVLAPLHSLAQALEHGLPDPGSL 240  
 DB 181 GPRVAPNPTGVDSAMKEEVQRLQSRVLDLLEKQLVLAPLHSLAQALEHGLPDPGSL 240

QY 241 VHSFQQLGRIDSLSEQISFLEEQLGSCCKDS 273  
 DB 241 VHSFQQLGRIDSLSEQISFLEEQLGSCCKDS 273

RESULT 6  
 AAB18673  
 ID AAB18673 standard; protein; 273 AA.  
 AC AAB18673;  
 XX  
 DT 22-JAN-2001 (first entry)  
 DE Amino acid sequence of a human a PRO213 polypeptide.  
 XX Growth arrest-specific gene 6 protein homologue; PRO320; PRO938; PRO1031;  
 KW PRO296; PRO213; PRO1449; angiogenesis; cardiovascularisation; PRO1330;  
 KW cardiovascular disorder; endothelial disorder; angiogenic disorder;  
 KW cancer; trauma; wound; arteriosclerosis; cardiac hypertrophy.  
 XX Homo sapiens.  
 XX  
 FT Key Location/Qualifiers  
 FT Peptide 1..19  
 FT Modified-site /note= "signal sequence"  
 FT Modified-site 19..25  
 FT Modified-site /note= "N-myristoylation site"  
 FT Modified-site 26..30  
 FT Modified-site /note= "amidation site"  
 FT Modified-site 78..84

FH Key Location/Qualifiers  
 FT Peptide 1..22  
 FT /label= Signal\_peptide  
 FT Protein 23..274  
 FT /label= TANGO\_125  
 FT Domain 107..134  
 FT /note= "EGF-like domain 1"  
 FT Domain 141..176  
 FT /note= "EGF-like domain 2"  
 XX  
 PN WO9954437-A2.  
 XX  
 XX 28-OCT-1999.  
 XX  
 XX 23-APR-1999; 99WO-US008900.  
 XX  
 PR 23-APR-1998; 98US-00065363.  
 PR 23-APR-1999; 99US-00298531.  
 XX  
 PA (MILL-) MILLENNIUM BIOTHERAPEUTICS INC.  
 XX  
 XX Holtzman DA;  
 XX  
 XX WPI; 2000-013240/01.  
 DR N-PSDB; AA237131.  
 XX  
 XX Novel polynucleotides and polypeptides used to modulate a variety of cellular processes.  
 PT  
 PT  
 PT  
 PS Claim 1; Fig 1; 120pp; English.  
 XX  
 CC This is the human TANGO 125 (T125) amino acid sequence. The T125 protein has two epidermal growth factor (EGF)-like domains at amino acids 107-134 and 141-176 and is predicted to have a molecular weight of approximately 141-176 and is predicted to have no transmembrane domains and appears to be a secreted protein. There are three alternatively spliced forms of the T125 gene: T125a and T125b and T125c (AA237132-237135). The sequences of all variants of T125 are used in the invention to create antibodies which selectively bind to T125. The T125 polypeptide is used to modulate a variety of cellular processes. It can be used to produce fusion proteins. The protein may also be used to produce antibodies, and to identify T125 antagonists and agonists. The T125 polynucleotides, polypeptides, homologues and antibodies can be used in screening assays; predictive medicine; and methods of treatment of T125 associated disorders. The T125 polynucleotides can be used to express the protein; to detect T125 mRNA; to detect genetic alterations in the T125 gene; in forensic biology; and as a source of primers and probes  
 CC  
 XX Sequence 273 AA;  
 SQ  
 Query Match 100.0%; Score 1505; DB 3; Length 273;  
 Best Local Similarity 100.0%; Pred. No. 1.3e-93;  
 Matches 273; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MRGSEVLLMVLAVGTEHAYRPGRRVCAVRAHGDVPVSEFVQRYQFLITCDGHR 60  
 Db 1 MRGSEVLLMVLAVGTEHAYRPGRRVCAVRAHGDVPVSEFVQRYQFLITCDGHR 60  
 QY 61 ACSTRTTYRTAYRSPGLAPRYPACPGWKTSGLPACGAAICOPPCRNCGSCVQP 120  
 Db 61 ACSTRTTYRTAYRSPGLAPRYPACPGWKTSGLPACGAAICOPPCRNCGSCVQP 120  
 QY 121 GRCPAGWGTGTCSDVDCEARRGCGPQRCINTAGSYWCOCWEGHLSADGTLCPVK 180  
 Db 121 GRCPAGWGTGTCSDVDCEARRGCGPQRCINTAGSYWCOCWEGHLSADGTLCPVK 180  
 QY 181 GPRVAPNTGVDNKAKEVQLQSRVDLLEKQLVLAPLHSLASQALEHGLPDGSL 240  
 Db 181 GPRVAPNTGVDNKAKEVQLQSRVDLLEKQLVLAPLHSLASQALEHGLPDGSL 240  
 QY 241 VHSFQQLGRIDSLSEQISFLEQLGSCCKDS 273  
 Db 241 VHSFQQLGRIDSLSEQISFLEQLGSCCKDS 273

## RESULT 4

AAB44326  
 ID AAB44326 standard; protein; 273 AA.  
 XX  
 AC AAB44326;  
 XX  
 DT 08-FEB-2001 (first entry)  
 XX  
 XX Human PRO1330 protein sequence SEQ ID NO:508.  
 XX  
 DE  
 XX  
 KW Human; secreted protein; transmembrane protein; PRO; EST; cytosstatic;  
 KW expressed sequence tag; detection; cancer.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200053756-A2.  
 XX  
 XX 14-SEP-2000.  
 XX  
 XX 18-FEB-2000; 2000WO-US004341.  
 XX  
 PR 08-MAR-1999; 99WO-US005028.  
 PR 12-MAR-1999; 99US-0123957P.  
 PR 29-MAR-1999; 99US-0126773P.  
 PR 21-APR-1999; 99US-0130232P.  
 PR 28-APR-1999; 99US-0131445P.  
 PR 14-MAY-1999; 99US-0134287P.  
 PR 23-JUN-1999; 99US-0141037P.  
 PR 28-JUL-1999; 99US-0145898P.  
 PR 29-OCT-1999; 99US-0162506P.  
 PR 30-NOV-1999; 99WO-US028313.  
 PR 02-DEC-1999; 99WO-US028551.  
 PR 16-DEC-1999; 99WO-US028565.  
 PR 30-DEC-1999; 99WO-US030095.  
 PR 30-DEC-1999; 99WO-US031243.  
 PR 30-DEC-1999; 99WO-US031274.  
 PR 03-JAN-2000; 2000WO-US000219.  
 PR 06-JAN-2000; 2000WO-US000277.  
 PR 06-JAN-2000; 2000WO-US000376.  
 XX  
 (GETH ) GENENTECH INC.  
 XX  
 XX Ashkenazi AJ, Baker XP, Botstein D, Deenoyers L, Eaton DL;  
 PI Ferrara N, Filvaroff E, Fong S, Gao W, Gerber H, Gerritsen ME;  
 PI Goddard A, Godowski PJ, Grimaldi CJ, Gurney AL, Hillan KJ;  
 PI Kijavini IJ, Kuo SS, Napier MA, Pan J, Paoni NF, Roy MA, Shelton DL;  
 PI Stewart TA, Tumas D, Williams PM, Wood WI;  
 XX  
 WPI; 2000-611443/58.  
 DR N-PSDB; AAC78586.  
 XX  
 PT Novel PRO polypeptides and polynucleotides used in detection methods, to target bioactive molecules to specific cells, and to modulate cellular activities.  
 PT  
 XX  
 PS Claim 12; Fig 215; 635pp; English.  
 XX  
 CC AAC78458 to AAC78599 represent polynucleotide and EST (expressed sequence tag) sequences which encode secreted or transmembrane PRO polypeptides. The PRO polynucleotides and polypeptides have cytosstatic activity. The polynucleotides and polypeptides can be used for detecting the presence of PRO polypeptides in samples, for linking bioactive molecules to cells and for modulating biological activities of cells, using the polypeptides for specific targeting. The polypeptide targeting, can be used to kill the target cells, e.g. for the treatment of cancers. The polypeptide pairs provide specific targeting of bioactive molecules to cells. AAC78600 to AAC78987 represent PCR primers and probes used in the isolation of the PRO polynucleotide sequences  
 CC  
 XX Sequence 273 AA;  
 SQ

20-MAR-1998; 98US-0078939P.  
 25-MAR-1998; 98US-0079294P.  
 26-MAR-1998; 98US-0079656P.  
 27-MAR-1998; 98US-0079663P.  
 27-MAR-1998; 98US-0079664P.  
 27-MAR-1998; 98US-0079669P.  
 27-MAR-1998; 98US-0079728P.  
 27-MAR-1998; 98US-0079786P.  
 30-MAR-1998; 98US-0079920P.  
 30-MAR-1998; 98US-0080105P.  
 31-MAR-1998; 98US-0080107P.  
 31-MAR-1998; 98US-0080165P.  
 31-MAR-1998; 98US-0080194P.  
 01-APR-1998; 98US-0080327P.  
 01-APR-1998; 98US-0080328P.  
 01-APR-1998; 98US-0080333P.  
 01-APR-1998; 98US-0080334P.  
 08-APR-1998; 98US-0081049P.  
 08-APR-1998; 98US-0081070P.  
 08-APR-1998; 98US-0081071P.  
 08-APR-1998; 98US-0081195P.  
 09-APR-1998; 98US-0081203P.  
 09-APR-1998; 98US-0081203P.  
 15-APR-1998; 98US-0081817P.  
 15-APR-1998; 98US-0081858P.  
 15-APR-1998; 98US-0081932P.  
 15-APR-1998; 98US-0081955P.  
 21-APR-1998; 98US-0082568P.  
 21-APR-1998; 98US-0082569P.  
 22-APR-1998; 98US-0082700P.  
 22-APR-1998; 98US-0082704P.  
 22-APR-1998; 98US-0082804P.  
 23-APR-1998; 98US-0082767P.  
 23-APR-1998; 98US-0082796P.  
 27-APR-1998; 98US-0083336P.  
 28-APR-1998; 98US-0083332P.  
 28-APR-1998; 98US-0083392P.  
 29-APR-1998; 98US-0083495P.  
 29-APR-1998; 98US-0083496P.  
 29-APR-1998; 98US-0083499P.  
 29-APR-1998; 98US-0083500P.  
 29-APR-1998; 98US-0083545P.  
 29-APR-1998; 98US-0083554P.  
 29-APR-1998; 98US-0083558P.  
 29-APR-1998; 98US-0083559P.  
 30-APR-1998; 98US-0083742P.  
 05-MAY-1998; 98US-0084366P.  
 06-MAY-1998; 98US-0084414P.  
 06-MAY-1998; 98US-0084441P.  
 07-MAY-1998; 98US-0084598P.  
 07-MAY-1998; 98US-0084600P.  
 07-MAY-1998; 98US-0084627P.  
 07-MAY-1998; 98US-0084637P.  
 07-MAY-1998; 98US-0084639P.  
 07-MAY-1998; 98US-0084640P.  
 07-MAY-1998; 98US-0084643P.  
 13-MAY-1998; 98US-0085323P.  
 13-MAY-1998; 98US-0085338P.  
 13-MAY-1998; 98US-0085339P.  
 15-MAY-1998; 98US-0085573P.  
 15-MAY-1998; 98US-0085579P.  
 15-MAY-1998; 98US-0085580P.  
 15-MAY-1998; 98US-0085582P.  
 15-MAY-1998; 98US-0085689P.  
 15-MAY-1998; 98US-0085697P.  
 15-MAY-1998; 98US-0085700P.  
 15-MAY-1998; 98US-0085704P.  
 18-MAY-1998; 98US-0086023P.  
 22-MAY-1998; 98US-0086392P.  
 22-MAY-1998; 98US-0086414P.  
 22-MAY-1998; 98US-0086430P.  
 22-MAY-1998; 98US-0086486P.

28-MAY-1998; 98US-0087098P.  
 28-MAY-1998; 98US-0087106P.  
 28-MAY-1998; 98US-0087208P.  
 30-JUL-1998; 98US-0094651P.  
 11-SEP-1998; 98US-0100038P.  
 XX (GETH) GENENTECH INC.  
 XX Wood WI, Goddard A, Gurney A, Yuan J, Baker KP, Chen J;  
 PI WPI: 1999-551358/46.  
 DR N-PSDB; ANZ34312.  
 XX  
 PT New secreted and transmembrane polypeptides and their polynucleotides,  
 PT useful for treating blood coagulation disorders, cancers and cellular  
 PT adhesion disorders.  
 XX  
 XX Claim 12; Fig 215; 530pp; English.  
 XX  
 CC The present invention describes secreted and transmembrane polypeptides  
 CC and their polynucleotides. The nucleotide sequences are useful as sources  
 CC of probes, primers, for chromosome mapping, and for generation of  
 CC antisense sequences. They can also be used to create transgenic animals.  
 CC The proteins can be used to treat a variety of diseases and disorders,  
 CC depending on their function. Diseases that may be treated include blood  
 CC coagulation disorders, cancers and cellular adhesion disorders. They may  
 CC also be used to raise antibodies. AA233891 to AA234338, and AA41685 to  
 CC AA41774 represent polynucleotide and polypeptide sequence given in the  
 CC exemplification of the present invention  
 XX  
 SQ Sequence 273 AA;  
 Query Match 100.0%; Score 1505; DB 2; Length 273;  
 Best Local Similarity 100.0%; Pred. No. 1.3e-93;  
 Matches 273; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MRGSEVLLMWLLVLAAGTGHAYRPPRRVCAVRAHGDVPSSEFVQRYVQPLTTCDGHR 60  
 DB 1 MRGSEVLLMWLLVLAAGTGHAYRPPRRVCAVRAHGDVPSSEFVQRYVQPLTTCDGHR 60  
 QY 61 ACSTYRTIYTRYSPLARPRYACCPGWRKTSGLPGAGAAICOPPCRNCGSCVQP 120  
 DB 61 ACSTYRTIYTRYSPLARPRYACCPGWRKTSGLPGAGAAICOPPCRNCGSCVQP 120  
 QY 121 GRCRCPAGWRGDTQSDVDECSARRGGCFQRCINTAGSYWCQCEGHSLSADGTLCPVKG 180  
 DB 121 GRCRCPAGWRGDTQSDVDECSARRGGCFQRCINTAGSYWCQCEGHSLSADGTLCPVKG 180  
 QY 181 GPPRVAPNPTGVDSAMKEEVQRLQSRVDLLEKQLVLPLHSLASQALEHGLPDGSL 240  
 DB 181 GPPRVAPNPTGVDSAMKEEVQRLQSRVDLLEKQLVLPLHSLASQALEHGLPDGSL 240  
 QY 241 VHSFQOLGRIDSLSEIQISFLEQLGSCCKKDS 273  
 DB 241 VHSFQOLGRIDSLSEIQISFLEQLGSCCKKDS 273  
 RESULT 3  
 AAY52137  
 ID AAY52137 standard; protein; 273 AA.  
 XX  
 AC AAY52137;  
 XX  
 DT 01-FEB-2000 (first entry)  
 XX  
 DE Human TANGO 125 (T125) amino acid sequence.  
 XX  
 KW TANGO 125; T125; alternative splice variant; EGF domain; antibody;  
 KW secreted protein; agonist; antagonist; predictive medicine; treatment;  
 XX forensic biology.  
 OS Homo sapiens.  
 XX



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 22, 2004, 12:41:14 ; Search time 61 Seconds  
(without alignments)  
1264.516 Million cell updates/sec

Title: US-09-978-191a-506  
Perfect score: 1505  
Sequence: 1 MRGSEVLLMWLLVAVGGT.....SEKISFLEQLGSCCKKDS 273

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_29Jan04.\*  
1: Geneseq1980s.\*  
2: Geneseq1990s.\*  
3: Geneseq2000s.\*  
4: Geneseq2001s.\*  
5: Geneseq2002s.\*  
6: Geneseq2003as.\*  
7: Geneseq2003bs.\*  
8: Geneseq2004s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1505	100.0	273	2	AA41769 Human PRO
2	1505	100.0	273	2	AA41770 Human PRO
3	1505	100.0	273	3	AA52137 Human TAN
4	1505	100.0	273	3	AA44326 Human PRO
5	1505	100.0	273	3	AA44325 Human PRO
6	1505	100.0	273	3	AA44325 Human PRO
7	1505	100.0	273	3	AA44325 Human PRO
8	1505	100.0	273	3	AA44325 Human PRO
9	1505	100.0	273	3	AA44325 Human PRO
10	1505	100.0	273	3	AA44325 Human PRO
11	1505	100.0	273	3	AA44325 Human PRO
12	1505	100.0	273	3	AA44325 Human PRO
13	1505	100.0	273	3	AA44325 Human PRO
14	1505	100.0	273	3	AA44325 Human PRO
15	1505	100.0	273	3	AA44325 Human PRO
16	1505	100.0	273	3	AA44325 Human PRO
17	1505	100.0	273	3	AA44325 Human PRO
18	1505	100.0	273	3	AA44325 Human PRO
19	1505	100.0	273	3	AA44325 Human PRO
20	1505	100.0	273	3	AA44325 Human PRO
21	1505	100.0	273	3	AA44325 Human PRO
22	1505	100.0	273	3	AA44325 Human PRO
23	1505	100.0	273	3	AA44325 Human PRO
24	1505	100.0	273	3	AA44325 Human PRO
25	1505	100.0	273	3	AA44325 Human PRO

26	1505	100.0	273	6	ADA25047	Ada25047	Novel	hum
27	1505	100.0	273	6	ADA25045	Ada25045	Novel	hum
28	1505	100.0	273	6	ABO19727	AbO19727	Novel	hum
29	1505	100.0	273	6	ABO19726	AbO19726	Novel	hum
30	1505	100.0	273	6	ADA12706	Ada12706	Human	sec
31	1505	100.0	273	6	ADA12708	Ada12708	Human	sec
32	1505	100.0	273	6	ABJ72444	AbJ72444	Human	PRO
33	1505	100.0	273	6	ABO34339	AbO34339	Human	sec
34	1505	100.0	273	6	ABO19617	AbO19617	Novel	hum
35	1505	100.0	273	6	ABO19618	AbO19618	Novel	hum
36	1505	100.0	273	7	ABJ72146	AbJ72146	Human	mem
37	1505	100.0	273	7	ADB83668	AdB83668	Novel	hum
38	1505	100.0	273	7	ADB80774	AdB80774	Novel	hum
39	1505	100.0	273	7	ADB73315	AdB73315	Novel	hum
40	1505	100.0	273	7	ADB78397	AdB78397	Novel	hum
41	1505	100.0	273	7	ADB85045	AdB85045	Human	PRO
42	1505	100.0	273	7	ADB78151	AdB78151	Novel	hum
43	1505	100.0	273	7	ADB74014	AdB74014	Human	PRO
44	1505	100.0	273	7	ADB74012	AdB74012	Human	PRO
45	1505	100.0	273	7	ADB87217	AdB87217	Human	PRO

## ALIGNMENTS

RESULT 1  
AA41769  
ID AA41769 standard; protein; 273 AA.

XX AC AA41769;

XX DT 07-DEC-1999 (first entry)

XX DE Human PRO213-1 protein sequence.

XX KW Human; PRO; EST; expressed sequence tag; PCR primer; hybridisation;

XX KM probe; blood coagulation disorder; cancer; cellular adhesion disorder;

XX KX secreted protein; transmembrane protein.

XX OS Homo sapiens.

XX PN WO9946281-A2.

XX PD 16-SEP-1999.

XX PF 08-MAR-1999; 99WO-US005028.

XX PR 10-MAR-1998; 98US-0077450P.

XX PR 11-MAR-1998; 98US-0077632P.

XX PR 11-MAR-1998; 98US-0077641P.

XX PR 12-MAR-1998; 98US-0077649P.

XX PR 13-MAR-1998; 98US-0077791P.

XX PR 17-MAR-1998; 98US-0078004P.

XX PR 20-MAR-1998; 98US-0078886P.

XX PR 20-MAR-1998; 98US-0078910P.

XX PR 20-MAR-1998; 98US-0078936P.

XX PR 25-MAR-1998; 98US-0079294P.

XX PR 26-MAR-1998; 98US-0079656P.

XX PR 27-MAR-1998; 98US-0079663P.

XX PR 27-MAR-1998; 98US-0079664P.

XX PR 27-MAR-1998; 98US-0079689P.

XX PR 27-MAR-1998; 98US-0079728P.

XX PR 27-MAR-1998; 98US-0079786P.

XX PR 30-MAR-1998; 98US-0079920P.

XX PR 31-MAR-1998; 98US-0080105P.

XX PR 31-MAR-1998; 98US-0080107P.

XX PR 31-MAR-1998; 98US-0080165P.

XX PR 31-MAR-1998; 98US-0080194P.

XX PR 01-APR-1998; 98US-0080327P.

XX PR 01-APR-1998; 98US-0080328P.





long-term force bearing structural support.  
-!- PNM: Forms intermolecular disulfide bonds either with other fibrillin-1 molecules or with other components of the microfibrils (By similarity).  
-!- SIMILARITY: Contains 47 EGF-like domains.  
-!- SIMILARITY: Contains 7 TGF-beta binding protein (TGFBP) domains.  
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
EMBL; L29454; AAA56840.1; -;  
EMBL; U22493; AAA64217.1; -;  
PIR; A55624; A55624.  
HSGP; P35555; LAPJ.  
MGD; MGI:95489; Fbml.  
InterPro; IPR00152; Asx\_hydroxyl\_S.  
InterPro; IPR001881; EGF\_Ca.  
InterPro; IPR006209; EGF\_Like.  
InterPro; IPR002212; Fibril-assoc.  
Pfam; PF00008; EGF; 45.  
Pfam; PF00683; TE; 9.  
SMART; SM0179; EGF\_CA; 42.  
PROSITE; PS00010; ASX\_HYDROXYL; 43.  
PROSITE; PS00022; EGF\_1; 2.  
PROSITE; PS01186; EGF\_2; 38.  
PROSITE; PS00026; EGF\_3; 45.  
PROSITE; PS01187; EGF\_CA; 43.  
Extracellular matrix; Calcium-binding; Glycoprotein; EGF-like domain; Repeat; Signal; Multigene family.  
SIGNAL 1 27  
CHAIN 28 2871  
FT DOMAIN 81 112 FIBRILLIN 1.  
FT DOMAIN 115 146 EGF-LIKE 1.  
FT DOMAIN 147 178 EGF-LIKE 2.  
FT DOMAIN 246 287 EGF-LIKE 3.  
FT DOMAIN 288 329 EGF-LIKE 4, CALCIUM-BINDING.  
FT DOMAIN 330 401 EGF-LIKE 5, CALCIUM-BINDING.  
FT DOMAIN 402 446 TGFBP 1.  
FT DOMAIN 449 489 PRO-RICH.  
FT DOMAIN 490 529 EGF-LIKE 6.  
FT DOMAIN 530 571 EGF-LIKE 7, CALCIUM-BINDING.  
FT DOMAIN 572 612 EGF-LIKE 8, CALCIUM-BINDING.  
FT DOMAIN 613 653 EGF-LIKE 9, CALCIUM-BINDING.  
FT DOMAIN 656 721 EGF-LIKE 10, CALCIUM-BINDING.  
FT DOMAIN 723 764 TGFBP 2.  
FT DOMAIN 765 806 EGF-LIKE 11, CALCIUM-BINDING.  
FT DOMAIN 807 846 EGF-LIKE 12, CALCIUM-BINDING.  
FT DOMAIN 910 951 EGF-LIKE 13, CALCIUM-BINDING.  
FT DOMAIN 952 1018 EGF-LIKE 14, CALCIUM-BINDING.  
FT DOMAIN 1028 1069 TGFBP 3.  
FT DOMAIN 1070 1112 EGF-LIKE 15, CALCIUM-BINDING.  
FT DOMAIN 1113 1154 EGF-LIKE 16, CALCIUM-BINDING.  
FT DOMAIN 1155 1196 EGF-LIKE 17, CALCIUM-BINDING.  
FT DOMAIN 1197 1237 EGF-LIKE 18, CALCIUM-BINDING.  
FT DOMAIN 1238 1279 EGF-LIKE 19, CALCIUM-BINDING.  
FT DOMAIN 1280 1321 EGF-LIKE 20, CALCIUM-BINDING.  
FT DOMAIN 1322 1362 EGF-LIKE 21, CALCIUM-BINDING.  
FT DOMAIN 1363 1403 EGF-LIKE 22, CALCIUM-BINDING.  
FT DOMAIN 1404 1445 EGF-LIKE 23, CALCIUM-BINDING.  
FT DOMAIN 1446 1486 EGF-LIKE 24, CALCIUM-BINDING.  
FT DOMAIN 1487 1527 EGF-LIKE 25, CALCIUM-BINDING.  
FT DOMAIN 1528 1569 EGF-LIKE 26, CALCIUM-BINDING.  
FT DOMAIN 1569 1607 TGFBP 4.  
FT DOMAIN 1608 1647 EGF-LIKE 27, CALCIUM-BINDING.  
FT DOMAIN 1648 1688 EGF-LIKE 28, CALCIUM-BINDING.  
FT DOMAIN 1689 1758 TGFBP 5.  
FT DOMAIN 1759 1807 EGF-LIKE 29, CALCIUM-BINDING.  
FT DOMAIN 1808 1848 EGF-LIKE 30, CALCIUM-BINDING.  
FT DOMAIN 1849 1890 EGF-LIKE 31, CALCIUM-BINDING.  
-----  
FT DOMAIN 1891 1930 EGF-LIKE 32, CALCIUM-BINDING.  
FT DOMAIN 1930 1972 EGF-LIKE 33, CALCIUM-BINDING.  
FT DOMAIN 1972 2012 EGF-LIKE 34, CALCIUM-BINDING.  
FT DOMAIN 2012 2054 EGF-LIKE 35, CALCIUM-BINDING.  
FT DOMAIN 2054 2121 TGFBP 6.  
FT DOMAIN 2121 2165 EGF-LIKE 36, CALCIUM-BINDING.  
FT DOMAIN 2165 2205 EGF-LIKE 37, CALCIUM-BINDING.  
FT DOMAIN 2205 2246 EGF-LIKE 38, CALCIUM-BINDING.  
FT DOMAIN 2246 2287 EGF-LIKE 39, CALCIUM-BINDING.  
FT DOMAIN 2287 2332 EGF-LIKE 40, CALCIUM-BINDING.  
FT DOMAIN 2332 2400 TGFBP 7.  
FT DOMAIN 2400 2443 EGF-LIKE 41, CALCIUM-BINDING.  
FT DOMAIN 2443 2484 EGF-LIKE 42, CALCIUM-BINDING.  
FT DOMAIN 2484 2523 EGF-LIKE 43, CALCIUM-BINDING.  
FT DOMAIN 2523 2566 EGF-LIKE 44, CALCIUM-BINDING.  
FT DOMAIN 2566 2606 EGF-LIKE 45, CALCIUM-BINDING.  
FT DOMAIN 2606 2647 EGF-LIKE 46, CALCIUM-BINDING.  
FT DOMAIN 2647 2687 EGF-LIKE 47, CALCIUM-BINDING.  
FT DISULFID 85 94 BY SIMILARITY.  
FT DISULFID 94 100 BY SIMILARITY.  
FT DISULFID 100 102 BY SIMILARITY.  
FT DISULFID 102 111 BY SIMILARITY.  
FT DISULFID 111 129 BY SIMILARITY.  
FT DISULFID 129 134 BY SIMILARITY.  
FT DISULFID 134 145 BY SIMILARITY.  
FT DISULFID 145 160 BY SIMILARITY.  
FT DISULFID 160 166 BY SIMILARITY.  
FT DISULFID 166 177 BY SIMILARITY.  
FT DISULFID 177 200 BY SIMILARITY.  
FT DISULFID 200 262 BY SIMILARITY.  
FT DISULFID 262 271 BY SIMILARITY.  
FT DISULFID 271 286 BY SIMILARITY.  
FT DISULFID 286 304 BY SIMILARITY.  
FT DISULFID 304 313 BY SIMILARITY.  
FT DISULFID 313 328 BY SIMILARITY.  
FT DISULFID 328 353 BY SIMILARITY.  
FT DISULFID 353 460 BY SIMILARITY.  
FT DISULFID 460 474 BY SIMILARITY.  
FT DISULFID 474 488 BY SIMILARITY.  
FT DISULFID 488 504 BY SIMILARITY.  
FT DISULFID 504 513 BY SIMILARITY.  
FT DISULFID 513 528 BY SIMILARITY.  
FT DISULFID 528 546 BY SIMILARITY.  
FT DISULFID 546 555 BY SIMILARITY.  
FT DISULFID 555 570 BY SIMILARITY.  
FT DISULFID 570 587 BY SIMILARITY.  
FT DISULFID 587 596 BY SIMILARITY.  
FT DISULFID 596 611 BY SIMILARITY.  
FT DISULFID 611 628 BY SIMILARITY.  
FT DISULFID 628 637 BY SIMILARITY.  
FT DISULFID 637 652 BY SIMILARITY.  
FT DISULFID 652 739 BY SIMILARITY.  
FT DISULFID 739 748 BY SIMILARITY.  
FT DISULFID 748 763 BY SIMILARITY.  
FT DISULFID 763 781 BY SIMILARITY.  
FT DISULFID 781 790 BY SIMILARITY.  
FT DISULFID 790 805 BY SIMILARITY.  
FT DISULFID 805 821 BY SIMILARITY.  
FT DISULFID 821 830 BY SIMILARITY.  
FT DISULFID 830 845 BY SIMILARITY.  
FT DISULFID 845 926 BY SIMILARITY.  
FT DISULFID 926 935 BY SIMILARITY.  
FT DISULFID 935 950 BY SIMILARITY.  
FT DISULFID 950 1044 BY SIMILARITY.  
FT DISULFID 1044 1053 BY SIMILARITY.  
FT DISULFID 1053 1066 BY SIMILARITY.  
FT DISULFID 1066 1086 BY SIMILARITY.  
FT DISULFID 1086 1095 BY SIMILARITY.  
FT DISULFID 1095 1111 BY SIMILARITY.  
FT DISULFID 1111 1129 BY SIMILARITY.  
FT DISULFID 1129 1138 BY SIMILARITY.  
FT DISULFID 1138 1153 BY SIMILARITY.  
FT DISULFID 1153 1171 BY SIMILARITY.  
FT DISULFID 1171 1180 BY SIMILARITY.  
FT DISULFID 1180 1195 BY SIMILARITY.  
FT DISULFID 1195 1212 BY SIMILARITY.

DR EMBL; X84819; CAA57709.1; JOINED.  
 DR EMBL; X84820; CAA57709.1; JOINED.  
 DR EMBL; X84821; CAA57709.1; JOINED.  
 DR EMBL; X84822; CAA57709.1; JOINED.  
 DR EMBL; X84823; CAA57709.1; JOINED.  
 DR EMBL; X84824; CAA57709.1; JOINED.  
 DR EMBL; X84825; CAA57709.1; JOINED.  
 DR EMBL; X84826; CAA57709.1; JOINED.  
 DR EMBL; X84827; CAA57709.1; JOINED.  
 DR EMBL; X84828; CAA57709.1; JOINED.  
 DR EMBL; X84829; CAA57709.1; JOINED.  
 DR EMBL; X84830; CAA57709.1; JOINED.  
 DR EMBL; X84831; CAA57709.1; JOINED.  
 DR EMBL; X84832; CAA57709.1; JOINED.  
 DR EMBL; X84833; CAA57709.1; JOINED.  
 DR EMBL; X84834; CAA57709.1; JOINED.  
 DR EMBL; X84835; CAA57709.1; JOINED.  
 DR EMBL; X84836; CAA57709.1; JOINED.  
 DR EMBL; X84837; CAA57709.1; JOINED.  
 DR EMBL; M27445; AAA57261.1; -.  
 DR PIR; A33322; MNRUND.  
 DR PDB; INDX; 29-DEC-99.  
 DR Genew; HGNC:7821; NID.  
 DR MIM; 131390; -.  
 DR InterPro; IPR000152; Asx hydroxyl\_s.  
 DR InterPro; IPR001881; EGF\_Ca.  
 DR InterPro; IPR006209; EGF\_Like.  
 DR InterPro; IPR006605; G2F.  
 DR InterPro; IPR009017; GFP like.  
 DR InterPro; IPR000033; Ldl receptor rep.  
 DR InterPro; IPR003886; Nidogen ext.  
 DR InterPro; IPR000716; Thyroglobulin\_1.  
 DR Pfam; PF00008; EGF\_6.  
 DR Pfam; PF00058; ldl\_recept\_b; 3.  
 DR Pfam; PF00086; thyroglobulin\_1; 1.  
 DR SMART; SM00179; EGF\_CA; 2.  
 DR SMART; SM00682; G2F; 1.  
 DR SMART; SM00135; LV; 5.  
 DR SMART; SM00539; NIDO; 1.  
 DR SMART; SM00211; TV; 1.  
 DR PROSITE; PS00010; ASX\_HYDROXYL; 3.  
 DR PROSITE; PS00022; EGF\_1; 1.  
 DR PROSITE; PS01186; EGF\_2; 5.  
 DR PROSITE; PS00026; EGF\_3; 5.  
 DR PROSITE; PS01187; EGF\_CA; 2.  
 DR PROSITE; PS00484; THYROGLOBULIN\_1; 1.  
 KW Basement membrane; Extracellular matrix; Glycoprotein; Sulfation;  
 KW Signal; Calcium-binding; Repeat; EGF-like domain; Cell adhesion;  
 KW 3D-structure.  
 FT SIGNAL 1 28  
 FT CHAIN 29 1247  
 FT DOMAIN 29 669  
 FT DOMAIN 670 917  
 FT DOMAIN 918 1247  
 FT DOMAIN 386 426  
 FT DOMAIN 668 709  
 FT DOMAIN 710 751  
 FT DOMAIN 758 801  
 FT DOMAIN 802 840  
 FT DOMAIN 872 919  
 FT DOMAIN 989 1030  
 FT DOMAIN 1032 1073  
 FT DOMAIN 1075 1118  
 FT DOMAIN 1124 1163  
 FT DOMAIN 1208 1244  
 FT MOD\_RES 289 289  
 FT MOD\_RES 296 296  
 FT DISULFID 672 695  
 FT DISULFID 679 695  
 FT DISULFID 697 708  
 FT DISULFID 714 727  
 FT DISULFID 721 736  
 FT DISULFID 738 750

FT DISULFID 762 777 BY SIMILARITY.  
 FT DISULFID 769 787 BY SIMILARITY.  
 FT DISULFID 789 800 BY SIMILARITY.  
 FT DISULFID 806 817 BY SIMILARITY.  
 FT DISULFID 811 826 BY SIMILARITY.  
 FT DISULFID 828 839 BY SIMILARITY.  
 FT DISULFID 1212 1233 BY SIMILARITY.  
 FT DISULFID 1219 1232 BY SIMILARITY.  
 FT DISULFID 1234 1243 BY SIMILARITY.  
 FT SITE 702 704 CELL ATTACHMENT SITE.  
 FT CARBOHYD 1137 1137 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CONFLICT 33 34 EL > SS (IN REF. 2).  
 FT CONFLICT 37 42 FGFQG -> SAPDR (IN REF. 2).  
 FT CONFLICT 1115 1115 H -> T (IN REF. 3).  
 SQ SEQUENCE 1247 AA; 136488 MW; 4681B5B3CEC1575B CRC64;  
 Query Match 11.5%; Score 173.5; DB 1; Length 1247;  
 Best Local Similarity 25.8%; Pred. No. 6e-06;  
 Matches 80; Conservative 23; Mismatches 98; Indels 109; Gaps 18;  
 QY 20 TEHAYRGRV-----CAVRAHG-----DPVSEFVQRYVQPLTTCGHRACSTYR 66  
 DB 681 TNAACRPGPRTQFTCECSIGFRGDRGRTCYDIDECSE-----QP--SVCSHTICNNH- 730  
 QY 67 TIYRTAYRSPGLAPRPRVACCPGKRTSGLEPGACGAALCOPP-----CRNG----- 114  
 DB 731 -----PGTFRCEVEG;QFSD--EGTCVAVDQRPINVCETGLHNCIDIPOR 774  
 QY 115 GSCVQPG-----RCRCPAGWRGD--TCQSDVDCSARAGGCPQRCINTAGSYWCQWEGHS 168  
 DB 775 AQCIYTGSGSYTCCLPGFSGDQACQ--DVDEQCPSCRHPDAFCYNTPGFTCCQKPGY- 832  
 QY 169 LSADGTLCPV-----KGGPRVAPNPTG----- 191  
 DB 833 -QGDGPRCPVGEVEKTRCQHERHIIAGAATDQRPPIPGGLFVPCDAGHYAPTCQHG 891  
 QY 192 -----VDSAMKEVQRLQSRVDLLEKLQLVLAFLH--SLASQALEHGLPPGSLLVHS 243  
 DB 892 STGYCWCVDGRDGR-EVEGTRTRPGMTTPCLSTVAPBIHQGPVPTAVIPLP-PGTHLL-- 947  
 QY 244 FQQLGRIDSL 253  
 DB 948 FAOTGKIERL 957  
 RESULT 15  
 FBN1\_MOUSE  
 ID\_FBN1\_MOUSE STANDARD; PRT; 2871 AA.  
 AC Q61554; Q60826;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Fibrillin 1 precursor.  
 GN FBN1 OR FBN-1.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=95130561; PubMed=7829516;  
 RA Yin W, Germiller J, Sanguinetti C, Smiley E, Pangillan T.,  
 RA Pereira L., Ramirez F, Bonadio J.,  
 RA "Primary structure and developmental expression of Fbn-1, the mouse  
 RT fibrillin gene.";  
 RL J. Biol. Chem. 270:1798-1806 (1995).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CD-1; TISSUE=Kidney;  
 RA Oka K., Kumar A., Wada J., Liu Z., Kanwar Y.S.,  
 RL Submitted (Apr-1995) to the EMBL/GenBank/DBJ databases.  
 CC -!- FUNCTION: Structural component of connective tissue microfibrils  
 CC that binds calcium. Fibrillin-1-containing microfibrils provide



QY 131 --GTCQSDVDCSARRGCCPQRCINTAGSWCQWGHSLA-DGTL-----CVPKGG 181  
DB 377 GRGEGACQDVDCALGRSPCAQCTNTDGSFHCSEGYVLAGEDGTQCDQDVDCVGGPG 436  
QY 182 P 182  
DB 437 P 437

RESULT 13  
NOTC DROME STANDARD; PRT: 2703 AA.  
AC P07207: 097458; P04154; Q9W4T8;  
DT 01-NOV-1985 (Rel. 03, Created)  
DT 28-FEB-2004 (Rel. 41, Last sequence update)  
DT 15-MAR-2004 (Rel. 43, Last annotation update)  
DE NEUROGENIC locus Notch protein precursor.  
GN N OR EG:140611.1 OR EG:163A10.2 OR CG3936.  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=7227;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Oregon-R; TISSUE=Embryo;  
RX MEDLINE=86079539; PubMed=3935325;  
RA Wharton K.A., Johansen K.M., Xu T., Artavanis-Tsakonas S.;  
RT "Nucleotide sequence from the neurogenic locus notch implies a gene  
RT product that shares homology with proteins containing EGF-like  
RT repeats.";  
RL Cell 43:567-581(1985).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Canton-S, and Oregon-R; TISSUE=Embryo;  
RX MEDLINE=87064624; PubMed=3097517;  
RA Kidd S., Kelley M.R., Young M.W.;  
RT "Sequence of the notch locus of Drosophila melanogaster: relationship  
RT of the encoded protein to mammalian clotting and growth factors.";  
RL Mol. Cell. Biol. 6:3094-3108(1986).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Berkeley;  
RX MEDLINE=20196006; PubMed=10731132;  
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,  
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,  
RA Abril J.P., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
RA Beeson K.Y., Benos P.V., Berhan P.P., Bhandari D., Bolshakov S.,  
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,  
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
RA Foster K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,  
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
RA Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,  
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,  
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,

RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
RA Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
RA Wang Z.-Y., Wassaman D.A., Weinstock G.M., Weissbach J.,  
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Smith H.O.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
RL "The genome sequence of Drosophila melanogaster.";  
RN Science 287:2185-2195(2000).  
RN [4]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Oregon-R;  
RX MEDLINE=20196011; PubMed=10731137;  
RA Benos P.V., Gatt M.K., Ashburner M., Murphy L., Harris D.,  
RA Barrell B.G., Ferraz C., Vidal S., Brun C., Demaillet J., Cadieu E.,  
RA Dreano S., Gloux S., Lelaure V., Mottier S., Galibert F., Borkova D.,  
RA Minana B., Kafatos F.C., Louis C., Siden-Kiamos I., Bolshakov S.,  
RA Papagiannakis G., Spanos L., Cox S., Madieno E., de Pablos B.,  
RA Modolell J., Peter A., Schoettler P., Werner M., Mourikoti F.,  
RA Beinart N., Dove G., Schaefer U., Jaekle H., Bucheton A.,  
RA Callister D.M., Campbell L.A., Darlamitsou A., Henderson M.S.,  
RA McMillan P.J., Salles C., Tait E.A., Valenti P., Saunders R.D.C.,  
RA Glover D.M.;  
RT "From sequence to chromosome: the tip of the X chromosome of D.  
RT melanogaster.";  
RL Science 287:2220-2222(2000).  
RN [5]  
RP SEQUENCE OF 2505-2611 FROM N.A.  
RX MEDLINE=85093329; PubMed=2981631;  
RA Wharton K.A., Fedvobnick B., Finnerty V.G., Artavanis-Tsakonas S.;  
RT "Opa: a novel family of transcribed repeats shared by the Notch locus  
RT and other developmentally regulated loci in D. melanogaster.";  
RL Cell 40:55-62(1985).  
RN [6]  
RP SEQUENCE OF 1-8 FROM N.A.  
RX MEDLINE=87257846; PubMed=3037327;  
RA Kelley M.R., Kidd S., Berg R.L., Young M.W.;  
RT "Restriction of P-element insertions at the Notch locus of Drosophila  
RT melanogaster.";  
RL Mol. Cell. Biol. 7:1545-1548(1987).  
RN [7]  
RP INTERACTION WITH DX, AND MUTANT SU42C.  
RX MEDLINE=94215489; PubMed=8162849;  
RA Diederich R.J., Matsuno K., Hing H., Artavanis-Tsakonas S.;  
RT "Cytosolic interaction between deltex and Notch ankyrin repeats  
RT implicates deltex in the Notch signaling pathway.";  
RL Development 120:473-481(1994).  
RN [8]  
RP INTERACTION WITH DX.  
RX MEDLINE=95401878; PubMed=7671895;  
RA Matsuno K., Diederich R.J., Go M.J., Blaumueller C.M.,  
RA Artavanis-Tsakonas S.;  
RT "Deltex acts as a positive regulator of Notch signaling through  
RT interactions with the Notch ankyrin repeats.";  
RL Development 121:2633-2644(1995).  
RN [9]  
RP S3 CLEAVAGE BY PSN.  
RX MEDLINE=99221487; PubMed=10206646;  
RA Struhl G., Greenwald I.;  
RT "Presenilin is required for activity and nuclear access of Notch in  
RT Drosophila.";  
RL Nature 398:522-525(1999).  
RN [10]  
RP S3 CLEAVAGE BY PSN.  
RX MEDLINE=99221498; PubMed=10206647;  
RA Ye Y., Lukinova N., Fortini M.E.;  
RT "Neurogenic phenotypes and altered Notch processing in Drosophila  
RT Presenilin mutants.";  
RL Nature 398:525-529(1999).  
RN [11]  
RP S2 CLEAVAGE BY KUZ.  
RX MEDLINE=21657146; PubMed=11799064;

RA Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,  
RA Rogers J.;  
RA "The DNA sequence and comparative analysis of human chromosome 20.";  
RL Nature 414:865-871(2001).  
[4]  
RN  
RP SEQUENCE FROM N.A.  
RC TISSUE=Leukocytes;  
RX MEDLINE=2238257; PubMed=12477932;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.W., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Usdin T.B., Toshyuki S., Carninci P., Prange C.,  
RA Raha S.S., Lequellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
RA Whiting K., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blackley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,  
RA Scherch A., Schein J.B., Jones S.J.M., Marra M.A.;  
RA "Generation and initial analysis of more than 15,000 full-length  
RT human and mouse cDNA sequences.";  
EL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
[5]  
RN  
RP CHARACTERIZATION.  
RX MEDLINE=21990337; PubMed=11994479;  
RA McGreal E.P., Ikewaki N., Akatsu H., Morgan B.P., Gasque P.;  
RA "Human ClqRp is identical with CD93 and the mN1-11 antigen but does  
RT not bind Clq.";  
RL J. Immunol. 168:5222-5232(2002).  
[6]  
RN  
RP O-GLYCOSYLATION.  
RX MEDLINE=99192777; PubMed=10092817;  
RA Nepomuceno R.R., Ruiz S., Park M., Tenner A.J.;  
RA "ClqRp is a heavily O-glycosylated cell surface protein involved in  
RT the regulation of phagocytic activity.";  
RL J. Immunol. 162:3583-3589(1999).  
CC  
CC -!- FUNCTION: Receptor (or element of a larger receptor complex) for  
CC Clq, mannose-binding lectin (MBL2) and pulmonary surfactant  
CC protein A (SPA). May mediate the enhancement of phagocytosis in  
CC monocytes and macrophages upon interaction with soluble defense  
CC collagens. May play a role in intercellular adhesion.  
CC  
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.  
CC  
CC -!- TISSUE SPECIFICITY: Highly expressed in endothelial cells,  
CC platelets, cells of myeloid origin, such as monocytes and  
CC neutrophils. Not expressed in cells of lymphoid origin.  
CC  
CC -!- PTM: N- and O-glycosylated.  
CC  
CC -!- SIMILARITY: Contains 1 C-type lectin family domain.  
CC  
CC -!- SIMILARITY: Contains 5 EGF-like domains.  
CC  
CC -!- CAUTION: Has been sometimes referred to as a collectin receptor.  
CC  
CC -!- CAUTION: According to Ref.5, Clq is not a ligand for ClqRA.  
CC  
CC -!- DATABASE: NAME=PRO; NOT=PROW 3:1-6(2001);  
CC WWW="http://www.ncbi.nlm.nih.gov/prow/guide/467246456.g.htm".  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC -----  
CC EMBL: U94333; AAB53110.1; --  
CC EMBL: AL118508; CAC00597.1; --  
CC EMBL: BC028075; AAB28075.1; --  
CC HSSP: P35555; 1EMW.  
CC Genew: HGNC:15855; ClQR1.  
CC MIN: 120577; --

DR GO: GO:0016021; C:integral to membrane; IC.  
DR GO: GO:0004872; F:receptor activity; NAS.  
DR GO: GO:0016337; P:cell-cell adhesion; IDA.  
DR GO: GO:0042116; P:macrophage activation; NAS.  
DR GO: GO:0006903; P:phagocytosis; NAS.  
DR InterPro: IPR001152; ASX\_hydroxyl\_S.  
DR InterPro: IPR001861; EGF\_Ca.  
DR InterPro: IPR006209; EGF-like.  
DR InterPro: IPR001304; Lectin\_C.  
DR Pfam: PF00008; EGF; 5.  
DR Pfam: PF00059; lectin\_c; 1.  
DR SMART: SM00034; CLECT; 1.  
DR SMART: SM00179; EGF\_Ca; 3.  
DR PROSITE: PS00010; ASX\_HYDROXYL; 3.  
DR PROSITE: PS00615; C-TYPE LECTIN\_1; FALSE NEG.  
DR PROSITE: PS00441; C-TYPE LECTIN\_2; 1.  
DR PROSITE: PS01186; EGF\_2; 3.  
DR PROSITE: PS01187; EGF\_Ca; 3.  
DR PROSITE: PS00026; EGF\_3; 3.  
KW Cell adhesion; Receptor; Repeat; Signal; Transmembrane;  
KW EGF-like domain; Lectin; Glycoprotein; Polymorphism.  
FT SIGNAL 1 21  
FT CHAIN 22 652  
FT DOMAIN 24 580  
FT TRANSMEM 581 601  
FT DOMAIN 602 652  
FT DOMAIN 32 174  
FT DOMAIN 260 301  
FT DOMAIN 302 344  
FT DOMAIN 345 384  
FT DOMAIN 385 426  
FT DOMAIN 427 468  
FT DOMAIN 594 601  
FT DISULFID 264 275  
FT DISULFID 211 285  
FT DISULFID 287 300  
FT DISULFID 306 317  
FT DISULFID 311 328  
FT DISULFID 330 343  
FT DISULFID 349 358  
FT DISULFID 354 367  
FT DISULFID 369 383  
FT DISULFID 389 400  
FT DISULFID 396 409  
FT DISULFID 411 425  
FT DISULFID 431 443  
FT DISULFID 439 452  
FT DISULFID 454 467  
FT CARBOHYD 325 325  
FT VARIANT 318 318  
FT  
FT CONFLICT 22 22 /FTID=VAR\_013573.  
FT CONFLICT 36 36 T -> V (IN REF. 1; AA SEQUENCE).  
FT CONFLICT 38 39 C -> T (IN REF. 1; AA SEQUENCE).  
FT CONFLICT 155 155 TA -> RI (IN REF. 1; AA SEQUENCE).  
FT CONFLICT 186 186 S -> N (IN REF. 1).  
FT CONFLICT 186 186 G -> A (IN REF. 1; AA SEQUENCE).  
FT CONFLICT 452 492 S -> A (IN REF. 1; AA SEQUENCE).  
FT CONFLICT 496 496 R -> Q (IN REF. 1; AA SEQUENCE).  
FT CONFLICT 504 504 R -> G (IN REF. 1; AA SEQUENCE).  
FT CONFLICT 541 541 P -> S (IN REF. 1).  
SQ SEQUENCE 652 AA; 68560 MW; EECA0FEAC55FCAC2 CRC64;  
Query Match 11.6%; Score 175; DB 1; Length 652;  
Best Local Similarity 33.1%; Pred. No. 2.3e-06;  
Matches 60; Conservative 10; Mismatches 57; Indels 54; Gaps 13;

QY 25 REGREV-----CAVRAHGDPVSESFQVRVYQPFLLTCDGHRACSTYRTIYRTAYRRSPG 78  
DB 288 RGFELLDDLYTCASR---NFCSSS-----PCRGATCV----- 318  
QY 79 LAPAPRYAC-CP-GWKSTSLGPGACGAAICQ-PPCRNGGSCVQ-PG--RCRCAGWR-- 130  
DB 319 LGPHGKNYTCRCPOGYQLDSQLDCVDDECQDSFCAQ--ECVNTPGFRCCEWGYBPG 376



Db 131 DDHCLQKGYIGTH-----CQPVCEGCLNGGRCVAPNRCACCTGTGTGQCERDRTGTP 185  
Qy 136 -----S 136  
Db 186 CFTVISNQMGQSLGIVCTYKQLCCATVGRANGHPCEMCPAHPCCRGPIPIRITGACQ 245  
Qy 137 DVDECSARRGCP-ORCINTAGSWCOCWEGHLSADGTLIC-----VPMKG--- 181  
Db 246 DVDECAIPGLCQCGNCINTVGSFECPCFAGHKLNEVSQKCEDIDECSTIPGICGEGECT 305  
Qy 182 -----PPRVAPNPTG 191  
Db 306 NTVSSYFCKCPGPGYTSFDG 325  
RESULT 11  
NTC3 HUMAN  
ID NTC3 HUMAN STANDARD; PROT: 2321 AA.  
AC Q9DMA7; OSUEB3; Q9UPL3; Q9Y6L8;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, last sequence update)  
DT 28-FEB-2003 (Rel. 41, last annotation update)  
DE Neurogenic locus notch homolog protein 3 precursor (Notch 3).  
GN NOTCH3.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_TaxID=9606;  
[1]  
RN RP MEDLINE FROM N.A.  
RX MEDLINE=9703728; PubMed=8878478;  
RA Joutel A., Corpechot C., Ducros A., Vahedi K., Chabriet H., Mouton P.,  
Alamowitch S., Domenga V., Cecillion M., Marechal E., Maciazek J.,  
Vayssiere C., Cruaud C., Cabanis E.-A., Ruchoux M.M., Weissenbach J.,  
Bach J.-F., Bousser M.-G., Tournier-Lasserre E.;  
RT "Notch3 mutations in CADASIL, a hereditary adult-onset condition  
causing stroke and dementia.";  
RL Nature 383:707-710(1996).  
[2]  
RN RP SEQUENCE FROM N.A.  
RA Gunel M., Artavanis-Tsakonas S.;  
RN Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.  
[3]  
RN RP SEQUENCE FROM N.A.  
RA Lamerdin J.E., McCreedy P.M., Skowronski E., Adamson A.W.,  
Burkhardt-Schultz K., Gordon L., Kyle A., Ramirez M., Stilwagen S.,  
Phan H., Velasco N., Ganes J., Danganan L., Poundstone P.,  
Christensen M., Georgescu A., Avila J., Liu S., Attix C., Andreise T.,  
Trankheim M., Amico-Keller G., Coefield J., Duarte S., Lucas S.,  
Bryce R., Thomas P., Quan G., Kronmiller B., Arellano A.,  
Montgomery M., Ow D., Nolan M., Trong S., Kobayashi A., Olsen A.S.,  
Carrano A.V.;  
RT "Sequence analysis of an 1.5 Mb olfactory receptor (OLPR) cluster in  
19p13.1";  
RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.  
[4]  
RN RP VARIANTS CADASIL TYR-49; CYS-71; CYS-90; CYS-110; CYS-133; CYS-141;  
ARG-146; CYS-153; CYS-169; CYS-171; CYS-182; ARG-185; SER-212;  
RP GLY-222; TYR-224; CYS-258; TYR-542; CYS-558; CYS-578; CYS-728;  
RP CYS-985; CYS-1006; CYS-1031; CYS-1231 AND ARG-1261, AND VARIANTS  
RP ARG-170; LEU-496; GLN-1133; MET-1183 AND ALA-2223.  
RX MEDLINE=98049753; PubMed=9388399;  
RA Joutel A., Vahedi K., Corpechot C., Troesch A., Chabriet H.,  
Vayssiere C., Cruaud C., Maciazek J., Weissenbach J., Bousser M.-G.,  
Bach J.-F., Tournier-Lasserre E.;  
RT "Strong clustering and stereotyped nature of Notch3 mutations in  
CADASIL patients.";  
RL Lancet 350:1511-1515(1997).  
[5]  
RN RP VARIANT CADASIL 114-GLY--PRO-120 DEL.  
RX MEDLINE=20264473; PubMed=10802807;  
RA Joutel A., Chabriet H., Vahedi K., Domenga V., Vayssiere C.,

RA Ruchoux M.M., Lucas C., Leys D., Bousser M.-G., Tournier-Lasserre E.;  
RT "Splice site mutation causing a seven amino acid Notch3 in-frame  
deletion in CADASIL.";  
RL Neurology 54:1874-1875(2000).  
[6]  
RN RP IDENTIFICATION OF LIGANDS.  
RX MEDLINE=99180765; PubMed=10079256;  
RA Gray G.B., Mann R.S., Mitsiadis E., Henrique D., Carcangiu M.-L.,  
Banks A., Leiman J., Ward D., Ish-Horowitz D., Artavanis-Tsakonas S.;  
RT "Human ligands of the Notch receptor.";  
RL Am. J. Pathol. 154:785-794(1999).  
CC -!- FUNCTION: Functions as a receptor for membrane-bound ligands  
Jagged1, Jagged2 and Delta1 to regulate cell-fate determination.  
Upon ligand activation through the released notch intracellular  
domain (NICD) it forms a transcriptional activator complex with  
RBP-J kappa and activates genes of the enhancer of split locus.  
CC Affects the implementation of differentiation, proliferation and  
apoptotic programs (By similarity).  
CC -!- SUBUNIT: Heterodimer of a C-terminal fragment N(TW) and a N-  
terminal fragment N(EC) which are probably linked by disulfide  
bonds (By similarity).  
CC -!- SUBCELLULAR LOCATION: Type I membrane protein. Following  
proteolytical processing NICD is translocated to the nucleus.  
CC -!- TISSUE SPECIFICITY: Ubiquitously expressed in fetal and adult  
tissues.  
CC -!- PTM: Synthesized in the endoplasmic reticulum as an inactive form  
which is proteolytically cleaved by a furin-like convertase in the  
trans-Golgi network before it reaches the plasma membrane to yield  
an active, ligand-accessible form. Cleavage results in a C-  
terminal fragment N(TW) and a N-terminal fragment N(EC). Following  
ligand binding, it is cleaved by TNF-alpha converting enzyme  
(TACE) to yield a membrane-associated intermediate fragment called  
notch extracellular truncation (NEXT). This fragment is then  
cleaved by presenilin dependent gamma-secretase to release a  
notch-derived peptide containing the intracellular domain (NICD)  
from the membrane (By similarity).  
CC -!- PTM: Phosphorylated (By similarity).  
CC -!- DISEASE: Defects in NOTCH3 are associated with cerebral autosomal  
dominant arteriopathy with subcortical infarcts and  
leukoencephalopathy (CADASIL) [MIM:125310]. CADASIL causes a type  
of stroke and dementia of which key features include recurrent  
subcortical ischemic events and vascular dementia.  
CC -!- SIMILARITY: Belongs to the NOTCH family.  
CC -!- SIMILARITY: Contains 34 EGF-like domains.  
CC -!- SIMILARITY: Contains 3 Lin/Notch repeats.  
CC -!- SIMILARITY: Contains 5 ANK repeats.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC EMBL: U97669; AAC91371.1;  
CC EMBL: AF058900; AAC14346.1;  
CC EMBL: AF058887; AAC14346.1; JOINED.  
CC EMBL: AF058882; AAC14346.1; JOINED.  
CC EMBL: AF058883; AAC14346.1; JOINED.  
CC EMBL: AF058884; AAC14346.1; JOINED.  
CC EMBL: AF058885; AAC14346.1; JOINED.  
CC EMBL: AF058886; AAC14346.1; JOINED.  
CC EMBL: AF058887; AAC14346.1; JOINED.  
CC EMBL: AF058888; AAC14346.1; JOINED.  
CC EMBL: AF058889; AAC14346.1; JOINED.  
CC EMBL: AF058890; AAC14346.1; JOINED.  
CC EMBL: AF058891; AAC14346.1; JOINED.  
CC EMBL: AF058892; AAC14346.1; JOINED.  
CC EMBL: AF058893; AAC14346.1; JOINED.  
CC EMBL: AF058894; AAC14346.1; JOINED.  
CC EMBL: AF058895; AAC14346.1; JOINED.  
CC EMBL: AF058896; AAC14346.1; JOINED.



[8]  
RN STRUCTURE BY NMR OF 2124-2205.  
RX MEDLINE=96222301; PubMed=8653794;  
RA Downing A.A., Knott V., Werner J.M., Cardy C.M., Campbell I.D.,  
RA Handford P.A.,  
RT "Solution structure of a pair of calcium-binding epidermal growth  
RT factor-like domains: implications for the Marfan syndrome and other  
RT genetic disorders.";  
RL Cell 85:597-605(1996).  
[9]  
RN REVIEW ON MFS VARIANTS.  
RX MEDLINE=96174615; PubMed=8594563;  
RA Colloff G., Beroud C., Soussi T., Junien C., Boileau C.;  
RT "Software and database for the analysis of mutations in the human  
RT FBN1 gene.";  
RL Nucleic Acids Res. 24:137-141(1996).  
[10]  
RN REVIEW ON MFS VARIANTS.  
RX MEDLINE=97169383; PubMed=9016526;  
RA Colloff-Beroud G., Beroud C., Ades L., Black C., Boxer M., Brock D.J.,  
RA Godfrey M., Hayward C., Karttunen L., Milewicz D., Peltonen L.,  
RA Richards R.I., Wang W., Junien C., Boileau C.;  
RT "Marfan Database (second edition): software and database for the  
RT analysis of mutations in the human FBN1 gene.";  
RL Nucleic Acids Res. 25:147-150(1997).  
[11]  
RN REVIEW ON VARIANTS.  
RX MEDLINE=98062175; PubMed=9401003;  
RA Hayward C., Brock D.J.H.;  
RT "Fibrillin-1 mutations in Marfan syndrome and other type-1  
RT fibrillinopathies.";  
RL Hum. Mutat. 10:415-423(1997).  
[12]  
RN VARIANTS MFS PRO-1137.  
RX MEDLINE=91304569; PubMed=1852208;  
RA Dietz H.C., Cutting G.R., Pyeritz R.E., Maslen C.L., Sakai L.Y.,  
RA Corson G.M., Puffenberger E.G., Hamosh A., Nanthakumar E.J.,  
RA Curristin S.M., Stetten G., Meyers D.A., Francomano C.A.;  
RT "Marfan syndrome caused by a recurrent de novo missense mutation in  
RT the fibrillin gene.";  
RL Nature 352:337-339(1991).  
[13]  
RN VARIANTS MFS SER-1249; ARG-1663; SER-2221 AND SER-2307.  
RX MEDLINE=93250834; PubMed=1301946;  
RA Dietz H.C., Saraiva J.M., Pyeritz R.E., Cutting G.R., Francomano C.A.;  
RT "Clustering of fibrillin (FBN1) missense mutations in Marfan syndrome  
RT patients at cysteine residues in EGF-like domains.";  
RL Hum. Mutat. 1:366-374(1992).  
[14]  
RN VARIANTS MFS SER-2307.  
RX MEDLINE=9223230; PubMed=1569206;  
RA Dietz H.C., Pyeritz R.E., Puffenberger E.G., Kendzior R.J., Jr.,  
RA Corson G.M., Maslen C.L., Sakai L.Y., Francomano C.A., Cutting G.R.;  
RT "Marfan phenotype variability in a family segregating a missense  
RT mutation in the epidermal growth factor-like motif of the fibrillin  
RT gene.";  
RL J. Clin. Invest. 89:1674-1680(1992).  
[15]  
RN VARIANTS MFS ILE-548 AND ALA-723.  
RX MEDLINE=94010946; PubMed=8406497;  
RA Dietz H.C., McIntosh I., Sakai L.Y., Corson G.M., Chalberg S.C.,  
RA Pyeritz R.E., Francomano C.A.;  
RT "Four novel FBN1 mutations: significance for mutant transcript level  
RT and EGF-like domain calcium binding in the pathogenesis of Marfan  
RT syndrome.";  
RL Genomics 17:468-475(1993).  
[16]  
RN VARIANTS MFS SER-2144.  
RX MEDLINE=93278402; PubMed=8504310;  
RA Hewett D.R., Lynch J.R., Smith R., Sykes B.C.;  
RT "A novel fibrillin mutation in the Marfan syndrome which could  
RT disrupt calcium binding of the epidermal growth factor-like module.";  
RL Hum. Mol. Genet. 2:475-477(1993).

[17]  
RN VARIANTS MFS ARG-862; TYR-1117; PRO-1137 AND PHE-1589, AND VARIANT  
RP ALA-1148.  
RX MEDLINE=94108431; PubMed=8281141;  
RA Tynan K., Comeau K., Pearson M., Wilgenbus P., Levitt D., Gasner C.,  
RA Berg M.A., Miller D.C., Francke U.;  
RT "Mutation screening of complete fibrillin-1 coding sequence: report  
RT of five new mutations, including two in 8-cysteine domains.";  
RL Hum. Mol. Genet. 2:1813-1821(1993).  
[18]  
RN VARIANTS MFS GLY-217 AND ARG-2627.  
RX MEDLINE=95067970; PubMed=7977366;  
RA Karttunen L., Raghunath M., Loeonqvist L., Peltonen L.;  
RT "A compound-heterozygous Marfan patient: two defective fibrillin  
RT alleles result in a lethal phenotype.";  
RL Am. J. Hum. Genet. 55:1083-1091(1994).  
[19]  
RN VARIANTS EL LYS-2447.  
RX MEDLINE=94245249; PubMed=8188302;  
RA Loeonqvist L., Child A., Kainulainen K., Davidson R., Puhakka L.,  
RA Peltonen L.;  
RT "A novel mutation of the fibrillin gene causing ectopia lentis.";  
RL Genomics 19:573-576(1994).  
[20]  
RN VARIANTS MFS CYS-627.  
RX MEDLINE=94272487; PubMed=8004112;  
RA Hayward C., Rae A.L., Porteous M.E.M., Logie L.J., Brock L.J.;  
RT "Two novel mutations and a neutral polymorphism in EGF-like domains  
RT of the fibrillin gene (FBN1): SSCP screening of exons 15-21 in Marfan  
RT syndrome patients.";  
RL Hum. Mol. Genet. 3:373-375(1994).  
[21]  
RN VARIANTS MFS CYS-122.  
RX MEDLINE=94314977; PubMed=8040326;  
RA Stahl-Hallengren C., Ukkonen T., Kainulainen K., Kristofersson U.,  
RA Saxne T., Tornqvist K., Peltonen L.;  
RT "An extra cysteine in one of the non-calcium-binding epidermal growth  
RT factor-like motifs of the FBN1 polypeptide is connected to a novel  
RT variant of Marfan syndrome.";  
RL J. Clin. Invest. 94:709-713(1994).  
[22]  
RN VARIANTS MFS TYR-1223.  
RX MEDLINE=94351692; PubMed=8071963;  
RA Hewett D.R., Lynch J.R., Child A., Sykes B.C.;  
RT "A new missense mutation of fibrillin in a patient with Marfan  
RT syndrome.";  
RL J. Med. Genet. 31:338-339(1994).  
[23]  
RN VARIANTS MFS HIS-1170.  
RX MEDLINE=95174777; PubMed=7870075;  
RA Hayward C., Porteous M.E.M., Brock D.J.H.;  
RT "A novel mutation in the fibrillin gene (FBN1) in familial  
RT arachnoidactyly.";  
RL Mol. Cell. Probes 8:325-327(1994).  
[24]  
RN VARIANTS MFS GLY-217; ASN-1023; ARG-1074; TYR-1242; ARG-1513;  
RP GLY-2127; TRP-2151; LYS-2447 AND ARG-2511.  
RX MEDLINE=94184368; PubMed=8136837;  
RA Kainulainen K., Karttunen L., Puhakka L., Sakai L., Peltonen L.;  
RT "Mutations in the fibrillin gene responsible for dominant ectopia  
RT lentis and neonatal Marfan syndrome.";  
RL Nat. Genet. 6:64-69(1994).  
[25]  
RN Query Match 11.7%; Score 176.5; DB 1; Length 2871;  
Best Local Similarity 22.7%; Pred. No. 6.5e-06;  
Matches 59; Conservative 19; Mismatches 79; Indels 103; Gaps 8;

OY 26 PGRVCAVRARHDPVSESFQVYQVPLTTCDEHR---ACSTVRYTYRTRSPGLAPA 82  
DB 75 PGGNQCLVPICHSCHGDFCSR---PNWCTCPGSIAPSCGS-PSIQHNCIRCMNGGSCS 130  
OY 83 RPRVACPGWKRKTSGLPGAGCAICQPPFCRNGSGSCVQPGRCRCPAGRGTCQ----- 135



Abbe T., Matsubashi S., Ting K.;  
 "Biochemical characterization and expression analysis of neural  
 thrombospondin-1-like proteins NELL1 and NELL2";  
 Biochem. Biophys. Res. Commun. 265:79-86(1999).  
 CC - SUBUNIT: Homotrimer. Binds to PKC beta-1.  
 CC - SUBCELLULAR LOCATION: Secreted.  
 CC - SIMILARITY: Contains 1 TSP N-terminal (TSPN) domain.  
 CC - SIMILARITY: Contains 5 WFCC domains.  
 CC - SIMILARITY: Contains 6 EGF-like domains.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----  
 EMBL; U48246; AAC72252.1; -  
 PIR; T10756; T10756.  
 DR HSP; P35555; 1EMN.  
 DR InterPro: IPR000152; Asx hydroxyl S.  
 DR InterPro: IPR008985; Cons like lec\_g1.  
 DR InterPro: IPR001881; EGF\_Ca.  
 DR InterPro: IPR006209; EGF\_Like.  
 DR InterPro: IPR001791; Laminin G.  
 DR InterPro: IPR003123; TSPN.  
 DR InterPro: IPR001007; VWF\_C.  
 DR Pfam; PF00008; EGF; 4.  
 DR Pfam; PF02210; TSPN; 1.  
 DR Pfam; PF00093; VWC; 2.  
 DR SMART; SM00179; EGF\_CA; 2.  
 DR SMART; SM00282; LamG; 1.  
 DR SMART; SM00210; TSPN; 1.  
 DR SMART; SM00214; VWC; 2.  
 DR PROSITE; PS00010; ASX HYDROXYL; 3.  
 DR PROSITE; PS00022; EGF\_1; 1.  
 DR PROSITE; PS01186; EGF\_2; 3.  
 DR PROSITE; PS00026; EGF\_3; 5.  
 DR PROSITE; PS01187; EGF\_CA; 3.  
 DR PROSITE; PS01208; VWF\_C; 2.  
 DR PROSITE; PS0184; VWF\_C; 2.  
 KW Glycoprotein; EGF-like domain; Repeat; Signal.  
 PT SIGNAL 1 16  
 FT CHAIN 17 810  
 FT DOMAIN 81 230  
 FT DOMAIN 271 332  
 FT DOMAIN 335 390  
 FT DOMAIN 391 433  
 FT DOMAIN 434 475  
 FT DOMAIN 476 516  
 FT DOMAIN 515 547  
 FT DOMAIN 549 595  
 FT DOMAIN 596 631  
 FT DOMAIN 632 687  
 FT DOMAIN 692 750  
 FT DOMAIN 752 807  
 FT DISULFID 395 407  
 FT DISULFID 401 416  
 FT DISULFID 418 432  
 FT DISULFID 438 451  
 FT DISULFID 445 460  
 FT DISULFID 462 474  
 FT DISULFID 480 493  
 FT DISULFID 487 502  
 FT DISULFID 504 515  
 FT DISULFID 519 529  
 FT DISULFID 523 535  
 FT DISULFID 537 546  
 FT DISULFID 553 566  
 FT DISULFID 560 575  
 FT DISULFID 577 594  
 FT DISULFID 600 613

FT DISULFID 607 622 BY SIMILARITY.  
 FT DISULFID 624 630 BY SIMILARITY.  
 FT CARBOHYD 40 40 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 53 53 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 83 83 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 224 224 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 294 294 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 372 372 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 511 511 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 562 562 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 609 609 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 708 708 N-LINKED (GLCNAC. .) (POTENTIAL).  
 SQ SEQUENCE 810 AA; 89212 MW; 46F09C466AF9A50B CRC64;  
 Query Match 12.0%; Score 180; DB 1; Length 810;  
 Best Local Similarity 36.5%; Pred. No. 1.2e-06;  
 Matches 35; Conservative 13; Mismatches 38; Indels 10; Gaps 3;  
 QY 88 CCGWKRTSLGACGAAICQPPCRNGGSCVQGRCRCPAGWRGDTQCSDVDECSARRGG 147  
 DB 504 CQPGYGVN-----GTICKAFCEGCRVGTCAVKNKVCPSGFTGSHCEKXIDECAGFYVE 559  
 QY 148 CP--QRCINTAGSYWCQWEGH---SLSADGTLV 177  
 DB 560 CHNYSRCVNLPGWYHCECRSGFHDGTYSLSGESCI 595  
 RESULT 9  
 ID\_NEL\_CHICK STANDARD; PRT; 816 AA.  
 AC Q90827;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 15-DEC-1998 (Rel. 37, Last annotation update)  
 DE NEL protein precursor (93 kDa protein).  
 GN NEL.  
 OS Gallus gallus (Chicken).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae;  
 OC Gallus.  
 OX NCBI\_TaxID=9031;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Embryo;  
 RX MEDLINE=95383734; PubMed=7655083;  
 RA Matsubashi S., Noji S., Koyama E., Myokai F., Ohuchi H.,  
 RA Taniguchi S., Hori K.;  
 RT "New gene, nel, encoding a M(r) 93 K protein with EGF-like repeats is  
 RT strongly expressed in neural tissues of early stage chick embryos.";  
 RL Dev. Dyn. 203:212-222(1995).  
 RN [2]  
 RP TSP N-TERMINAL DOMAIN.  
 RX MEDLINE=98153258; PubMed=9480764;  
 RA Beckmann G., Hanke J., Bork P., Reich J.;  
 RT "Merging extracellular domains: fold prediction for laminin G-like  
 RT and amino-terminal thrombospondin-like modules based on homology to  
 RT pentraxins.";  
 RL J. Mol. Biol. 275:725-730(1998).  
 CC - TISSUE SPECIFICITY: Strongly expressed in early embryonic neural  
 CC tissues (brain, spinal cord, dorsal root ganglia); less in other  
 CC tissues such as cells around cartilage, myocardium, lung  
 CC mesenchymal cells, and liver. After hatching expression is  
 CC restricted to neural tissues including retina.  
 CC - SIMILARITY: Contains 1 TSP N-terminal (TSPN) domain.  
 CC - SIMILARITY: Contains 5 WFCC domains.  
 CC - SIMILARITY: Contains 6 EGF-like domains.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC

QY 148 C--P--CINTRAGSWCQWEGH-----SLSDGTLGV 177  
DB 560 CHNHSRCVNLPGWYHCECRSGFDHDDGYSLSGESCI 595

RESULT 7  
NEL1 RAT  
ID NEL1 RAT STANDARD; PRT; 816 AA.  
AC Q62918;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Protein kinase C-binding protein NELL1 precursor (NEL-like protein 2).  
GN NELL2 OR NEL1  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
SEQUENCE FROM N.A.

RC STRAIN-Sprague-Dawley; TISSUE=Brain;  
RX MEDLINE=20017976; PubMed=10548494;  
RA Kuroda S., Oyasu M., Kawakami M., Kanayama N., Tanizawa K., Saito N.,  
Abe T., Matsushashi S., Ting K.;  
RT "Biochemical characterization and expression analysis of neural  
thrombospondin-1-like proteins NELL1 and NELL2";  
RL Biochem. Biophys. Res. Commun. 265:79-86(1999).  
RN [2]  
TSP N-TERMINAL DOMAIN.  
RX MEDLINE=98153258; PubMed=9480764;  
RA Beckmann G., Hanke J., Bork P., Reich J.;  
RT "Merging extracellular domains: fold prediction for laminin G-like  
and amino-terminal thrombospondin-like modules based on homology to  
pentraxins";  
RL J. Mol. Biol. 275:725-730(1998).

CC -!- SUBUNIT: Homotrimer. Binds to PKC beta-1.  
CC -!- SUBCELLULAR LOCATION: Secreted.  
CC -!- SIMILARITY: Contains 1 TSP N-terminal (TSPN) domain.  
CC -!- SIMILARITY: Contains 5 WFC domains.  
CC -!- SIMILARITY: Contains 6 EGF-like domains.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; U48245; AAC72245.1; -;  
DR HSP; P00740; IEDM.  
DR InterPro; IPR000152; Asx\_hydroxyl\_S.  
DR InterPro; IPR008985; ConA\_like\_lec\_gl.  
DR InterPro; IPR001881; EGF\_Ca.  
DR InterPro; IPR006209; EGF\_like.  
DR InterPro; IPR001791; Laminin\_G.  
DR InterPro; IPR003129; TSPN.  
DR InterPro; IPR001007; WFC\_C.  
DR Pfam; PF00008; EGF; 4.  
DR Pfam; PF02210; TSPN; 1.  
DR Pfam; PF00093; WFC; 3.  
DR SMART; SM00179; EGF\_CA; 3.  
DR SMART; SM00282; LamG; 1.  
DR SMART; SM00210; TSPN; 1.  
DR SMART; SM00214; WVC; 3.  
DR PROSITE; PS00010; ASX\_HYDROXYL; 3.  
DR PROSITE; PS00022; EGF\_1; 1.  
DR PROSITE; PS01186; EGF\_2; 4.  
DR PROSITE; PS00026; EGF\_3; 6.  
DR PROSITE; PS01187; EGF\_CA; 3.  
DR PROSITE; PS01208; WFC\_1; 2.  
DR PROSITE; PS0184; WFC\_2; 3.

KW Glycoprotein; EGF-like domain; Repeat; Signal.  
FT SIGNAL 1 24  
FT CHAIN 25 816  
FT DOMAIN 30 258  
FT DOMAIN 272 331  
FT DOMAIN 332 396  
FT DOMAIN 397 439  
FT DOMAIN 440 481  
FT DOMAIN 482 522  
FT DOMAIN 521 553  
FT DOMAIN 555 601  
FT DOMAIN 602 637  
FT DOMAIN 638 693  
FT DOMAIN 698 756  
FT DOMAIN 758 813  
FT DISULFID 401 413  
FT DISULFID 407 422  
FT DISULFID 424 438  
FT DISULFID 444 457  
FT DISULFID 451 466  
FT DISULFID 468 480  
FT DISULFID 486 499  
FT DISULFID 493 508  
FT DISULFID 510 521  
FT DISULFID 525 535  
FT DISULFID 529 541  
FT DISULFID 543 552  
FT DISULFID 559 572  
FT DISULFID 566 581  
FT DISULFID 583 600  
FT DISULFID 606 619  
FT DISULFID 613 628  
FT DISULFID 630 636  
FT CARBOHYD 53 53  
FT CARBOHYD 225 225  
FT CARBOHYD 293 293  
FT CARBOHYD 298 298  
FT CARBOHYD 517 517  
FT CARBOHYD 615 615  
FT CARBOHYD 635 635  
SQ SEQUENCE 816 AA; 90952 MW; A999F7678060D6B CRC64;  
Query Match 12.0%; Score 180.5; DB 1; Length 816;  
Best Local Similarity 36.8%; Pred. No. 1.1e-06;  
Matches 35; Conservative 14; Mismatches 37; Indels 9; Gaps 3;  
QY 88 CPGWKRISGLPGACGAALCQPPCRNGSCVOPGRCRCRCPAGWGTCCQSDVDECSARRGG 147  
DB 510 CKPGYTGK---GTTCKAFCKDCKNGGACIAANVCACPCQGTGSPCTDIDECSEGFVQ 565  
QY 148 CPOR--CINTAGSYWCQWEGHSLSDGTLGVKPKG 180  
DB 566 CDSRANCLNLPWYHCECRDGYH---DNGMFAPGG 597

RESULT 8  
NEL1 RAT  
ID NEL1 RAT STANDARD; PRT; 810 AA.  
AC Q62919;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Protein kinase C-binding protein NELL1 precursor (NEL-like protein 1).  
GN NELL1.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
SEQUENCE FROM N.A.  
RC STRAIN-Sprague-Dawley; TISSUE=Brain;  
RX MEDLINE=20017976; PubMed=10548494;  
RX Kuroda S., Oyasu M., Kawakami M., Kanayama N., Tanizawa K., Saito N.,

```
FT CARBOHYD 635 635 N-LINKED (GLCNAC... ) (POTENTIAL).
SQ SEQUENCE 816 AA; 91346 MW; 893708987DCTA324 CRC64;

Query Match
Best Local Similarity 12.1%; Score 182.5; DB 1; Length 816;
Matches 36; Conservative 13; Mismatches 37; Indels 9; Gaps 3;

QY 88 CPGWKRISGLGACGALICQPCNGGSCVQPGRCPCAGWRGDTQSDVDECSARRGG 147
DB 510 CKPGYTGNN---GTTCKAFKDCGRNGGACIAANVACPGQFTGPGSCETDIDECSDGFVQ 565

QY 148 CPQR--CINTAGSYWCQWEGHSLGADGTLCPVKG 180
DB 566 CDSRANCINLPGWYHCECHDGX---DNGMRPSPG 597

RESULT 6
NEL1 HUMAN
ID NEL1 HUMAN STANDARD; PRT; 810 AA.
AC Q92832; Q9V472;
DT 01-NOV-1997 (Rel. 35, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Protein kinase C-binding protein NELL1 precursor (NEL-like protein 1)
DE (Nel-related protein 1).
GN NELL1 OR NR1P1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=97131504; PubMed=8975702;
RA Watanabe T.K., Katagiri T., Suzuki M., Shimizu F., Fujiwara T.,
RA Kanenoto N., Nakamura Y., Hirai Y., Maekawa H., Takahashi E.;
RT "Cloning and characterization of two novel human cDNAs (NELL1 and
RL NELL2) encoding proteins with six EGF-like repeats.";
RL Genomics 38:273-276(1996).
RN [2]
SQ SEQUENCE OF 383-810 FROM N.A.
RA Ting K., Vastardis H., Mulliken J.B., Bertolami C., Wen Z.,
RA Young M., Tieu A., Kwong E.;
RT "Nel homolog gene expression in craniofacial anomalies.";
RL Submitted (Sep-1996) to the EMBL/GenBank/DBJ databases.
CC -1- SUBUNIT: Homotrimer. Binds to PKC beta-1 (By similarity).
CC -1- SUBCELLULAR LOCATION: Secreted (By similarity).
CC -1- DISEASE: Expressed in craniofacial anomalies.
CC -1- SIMILARITY: Contains 1 TSP N-terminal (TSPN) domain.
CC -1- SIMILARITY: Contains 5 WFC domains.
CC -1- SIMILARITY: Contains 6 EGF-like domains.
CC -1- CAUTION: Ref.2 sequence differs from that shown due to frameshifts
in positions 427 and 771.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; D83017; BAA11680.1; --
EMBL; U57523; AAB05946.1; ALT_FRAME.
HSP; P07204; IADX.
Genew; HGNC:7750; NELL1.
MIM; 602319; --
GO; GO:0007399; P:neurogenesis; TAS.
InterPro; IPR000152; Asx hydroxylase.
InterPro; IPR008985; ConA like lec_g1.
InterPro; IPR001881; EGF Ca.
InterPro; IPR006209; EGF-like.
InterPro; IPR001791; Laminin_G.
```

```
DR InterPro; IPR003129; TSPN.
DR InterPro; IPR001007; WFC_C.
DR Pfam; PF00008; EGF; 4.
DR Pfam; PF02210; TSPN; 1.
DR Pfam; PF00093; WFC; 2.
DR SMART; SM00179; EGF CA; 2.
DR SMART; SM00282; LamG; 1.
DR SMART; SM00210; TSPN; 1.
DR SMART; SM00214; WFC; 3.
DR PROSITE; PS00010; ASX_HYDROXYL; 3.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; 3.
DR PROSITE; PS00026; EGF_3; 5.
DR PROSITE; PS01187; EGF CA; 3.
DR PROSITE; PS01208; WFC 1; 2.
DR PROSITE; PS01084; WFC 2; 2.
KW Glycoprotein; EGF-like domain; Repeat; Signal.
FT SIGNAL 1 16
FT CHAIN 17 810
FT DOMAIN 81 230
FT DOMAIN 271 332
FT DOMAIN 335 390
FT DOMAIN 391 433
FT DOMAIN 434 475
FT DOMAIN 476 516
FT DOMAIN 515 547
FT DOMAIN 549 595
FT DOMAIN 631 631
FT DOMAIN 632 687
FT DOMAIN 692 750
FT DOMAIN 752 807
FT DISULFID 395 407
FT DISULFID 401 416
FT DISULFID 418 432
FT DISULFID 438 451
FT DISULFID 445 460
FT DISULFID 462 474
FT DISULFID 480 493
FT DISULFID 487 502
FT DISULFID 504 515
FT DISULFID 519 529
FT DISULFID 523 535
FT DISULFID 537 546
FT DISULFID 553 566
FT DISULFID 563 575
FT DISULFID 577 594
FT DISULFID 600 613
FT DISULFID 607 622
FT DISULFID 624 630
FT CARBOHYD 40 40
FT CARBOHYD 53 53
FT CARBOHYD 83 83
FT CARBOHYD 224 224
FT CARBOHYD 294 294
FT CARBOHYD 372 372
FT CARBOHYD 511 511
FT CARBOHYD 562 562
FT CARBOHYD 609 609
FT CARBOHYD 708 708
FT CARBOHYD 732 732
FT CARBOHYD 758 758
FT CONFLICT 383 383
FT CONFLICT 573 573
FT CONFLICT 626 626
SQ SEQUENCE 810 AA; 89606 MW; 549465EA3F7AEED0 CRC64;

Query Match 12.1%; Score 182; DB 1; Length 810;
Best Local Similarity 37.5%; Pred No. 8.1e-07;
Matches 36; Conservative 12; Mismatches 38; Indels 10; Gaps 3;

QY 88 CPGWKRISGLGACGALICQPCNGGSCVQPGRCPCAGWRGDTQSDVDECSARRGG 147
DB 504 CKPGYVGN---GTTCAFCBEGCRYGTCVAPNKCVCPSGFTGSHCKDIDECSEGIIE 559
```



```

568 609  EGF-LIKE 8,  CALCIUM-BINDING.
PT DOMAIN 568 609  EGF-LIKE 9,  CALCIUM-BINDING.
PT DOMAIN 610 650  EGF-LIKE 10,  CALCIUM-BINDING.
PT REPEAT 651 691  TGFBB 2.
PT REPEAT 692 760  EGF-LIKE 11,  CALCIUM-BINDING.
PT DOMAIN 761 802  EGF-LIKE 12,  CALCIUM-BINDING.
PT DOMAIN 803 844  EGF-LIKE 13,  CALCIUM-BINDING.
PT DOMAIN 845 883  EGF-LIKE 14,  CALCIUM-BINDING.
PT DOMAIN 884 924  TGFBB 3.
PT DOMAIN 925 965  EGF-LIKE 15,  CALCIUM-BINDING.
PT REPEAT 966 1006  EGF-LIKE 16,  CALCIUM-BINDING.
PT DOMAIN 1007 1107  EGF-LIKE 17,  CALCIUM-BINDING.
PT DOMAIN 1108 1151  EGF-LIKE 18,  CALCIUM-BINDING.
PT DOMAIN 1152 1234  EGF-LIKE 19,  CALCIUM-BINDING.
PT DOMAIN 1235 1275  EGF-LIKE 20,  CALCIUM-BINDING.
PT DOMAIN 1276 1317  EGF-LIKE 21,  CALCIUM-BINDING.
PT DOMAIN 1318 1359  EGF-LIKE 22,  CALCIUM-BINDING.
PT DOMAIN 1360 1401  EGF-LIKE 23,  CALCIUM-BINDING.
PT DOMAIN 1402 1441  EGF-LIKE 24,  CALCIUM-BINDING.
PT DOMAIN 1442 1483  EGF-LIKE 25,  CALCIUM-BINDING.
PT DOMAIN 1484 1525  EGF-LIKE 26,  CALCIUM-BINDING.
PT DOMAIN 1526 1566  TGFBB 4.
PT REPEAT 1567 1642  EGF-LIKE 27,  CALCIUM-BINDING.
PT DOMAIN 1643 1684  EGF-LIKE 28,  CALCIUM-BINDING.
PT REPEAT 1685 1726  TGFBB 5.
PT DOMAIN 1727 1800  EGF-LIKE 29,  CALCIUM-BINDING.
PT DOMAIN 1801 1842  EGF-LIKE 30,  CALCIUM-BINDING.
PT DOMAIN 1843 1884  EGF-LIKE 31,  CALCIUM-BINDING.
PT DOMAIN 1885 1926  EGF-LIKE 32,  CALCIUM-BINDING.
PT DOMAIN 1927 1965  EGF-LIKE 33,  CALCIUM-BINDING.
PT DOMAIN 1966 2008  EGF-LIKE 34,  CALCIUM-BINDING.
PT DOMAIN 2009 2048  EGF-LIKE 35,  CALCIUM-BINDING.
PT DOMAIN 2049 2090  TGFBB 6.
PT REPEAT 2091 2163  EGF-LIKE 36,  CALCIUM-BINDING.
PT DOMAIN 2164 2205  EGF-LIKE 37,  CALCIUM-BINDING.
PT DOMAIN 2206 2245  EGF-LIKE 38,  CALCIUM-BINDING.
PT DOMAIN 2246 2286  EGF-LIKE 39,  CALCIUM-BINDING.
PT DOMAIN 2287 2330  EGF-LIKE 40,  CALCIUM-BINDING.
PT DOMAIN 2331 2372  TGFBB 7.
PT REPEAT 2373 2441  EGF-LIKE 41,  CALCIUM-BINDING.
PT DOMAIN 2442 2483  EGF-LIKE 42,  CALCIUM-BINDING.
PT DOMAIN 2484 2524  EGF-LIKE 43,  CALCIUM-BINDING.
PT DOMAIN 2525 2563  EGF-LIKE 44,  CALCIUM-BINDING.
PT DOMAIN 2564 2606  EGF-LIKE 45,  CALCIUM-BINDING.
PT DOMAIN 2607 2646  EGF-LIKE 46,  CALCIUM-BINDING.
PT DOMAIN 2647 2687  EGF-LIKE 47,  CALCIUM-BINDING.
PT DOMAIN 2688 2727  BY SIMILARITY.
PT DISULFID 115 124  BY SIMILARITY.
PT DISULFID 119 130  BY SIMILARITY.
PT DISULFID 132 141  BY SIMILARITY.
PT DISULFID 149 159  BY SIMILARITY.
PT DISULFID 153 164  BY SIMILARITY.
PT DISULFID 166 175  BY SIMILARITY.
PT DISULFID 180 190  BY SIMILARITY.
PT DISULFID 184 196  BY SIMILARITY.
PT DISULFID 198 207  BY SIMILARITY.
PT DISULFID 280 292  BY SIMILARITY.
PT DISULFID 287 301  BY SIMILARITY.
PT DISULFID 303 316  BY SIMILARITY.
PT DISULFID 322 334  BY SIMILARITY.
PT DISULFID 329 343  BY SIMILARITY.
PT DISULFID 345 358  BY SIMILARITY.
PT DISULFID 491 503  BY SIMILARITY.
PT DISULFID 498 512  BY SIMILARITY.
PT DISULFID 514 526  BY SIMILARITY.
PT DISULFID 532 542  BY SIMILARITY.
PT DISULFID 537 551  BY SIMILARITY.
PT DISULFID 553 566  BY SIMILARITY.
PT DISULFID 572 584  BY SIMILARITY.
PT DISULFID 579 593  BY SIMILARITY.
PT DISULFID 595 608  BY SIMILARITY.
PT DISULFID 614 625  BY SIMILARITY.
PT DISULFID 620 634  BY SIMILARITY.
PT DISULFID 636 649  BY SIMILARITY.

655 666  DISULFID 655 666  BY SIMILARITY.
FT DISULFID 661 675  BY SIMILARITY.
FT DISULFID 677 690  BY SIMILARITY.
FT DISULFID 765 777  BY SIMILARITY.
FT DISULFID 772 786  BY SIMILARITY.
FT DISULFID 788 801  BY SIMILARITY.
FT DISULFID 807 819  BY SIMILARITY.
FT DISULFID 814 828  BY SIMILARITY.
FT DISULFID 830 843  BY SIMILARITY.
FT DISULFID 849 859  BY SIMILARITY.
FT DISULFID 854 868  BY SIMILARITY.
FT DISULFID 870 883  BY SIMILARITY.
FT DISULFID 952 964  BY SIMILARITY.
FT DISULFID 959 973  BY SIMILARITY.
FT DISULFID 975 988  BY SIMILARITY.
FT DISULFID 1070 1082  BY SIMILARITY.
FT DISULFID 1077 1091  BY SIMILARITY.
FT DISULFID 1093 1106  BY SIMILARITY.
FT DISULFID 1112 1124  BY SIMILARITY.
FT DISULFID 1119 1133  BY SIMILARITY.
FT DISULFID 1135 1149  BY SIMILARITY.
FT DISULFID 1155 1167  BY SIMILARITY.
FT DISULFID 1162 1176  BY SIMILARITY.
FT DISULFID 1178 1191  BY SIMILARITY.
FT DISULFID 1197 1209  BY SIMILARITY.
FT DISULFID 1204 1218  BY SIMILARITY.
FT DISULFID 1220 1233  BY SIMILARITY.
FT DISULFID 1239 1250  BY SIMILARITY.
FT DISULFID 1246 1259  BY SIMILARITY.
FT DISULFID 1261 1274  BY SIMILARITY.
FT DISULFID 1280 1292  BY SIMILARITY.
FT DISULFID 1287 1301  BY SIMILARITY.
FT DISULFID 1303 1316  BY SIMILARITY.
FT DISULFID 1322 1334  BY SIMILARITY.
FT DISULFID 1329 1343  BY SIMILARITY.
FT DISULFID 1345 1358  BY SIMILARITY.
FT DISULFID 1364 1377  BY SIMILARITY.
FT DISULFID 1371 1385  BY SIMILARITY.
FT DISULFID 1388 1399  BY SIMILARITY.
FT DISULFID 1405 1418  BY SIMILARITY.
FT DISULFID 1412 1427  BY SIMILARITY.
FT DISULFID 1429 1440  BY SIMILARITY.
FT DISULFID 1446 1458  BY SIMILARITY.
FT DISULFID 1453 1467  BY SIMILARITY.
FT DISULFID 1469 1482  BY SIMILARITY.
FT DISULFID 1488 1499  BY SIMILARITY.
FT DISULFID 1494 1508  BY SIMILARITY.
FT DISULFID 1510 1523  BY SIMILARITY.
FT DISULFID 1529 1540  BY SIMILARITY.
FT DISULFID 1535 1549  BY SIMILARITY.
FT DISULFID 1551 1564  BY SIMILARITY.
FT DISULFID 1567 1579  BY SIMILARITY.
FT DISULFID 1654 1668  BY SIMILARITY.
FT DISULFID 1670 1683  BY SIMILARITY.
FT DISULFID 1689 1701  BY SIMILARITY.

Query Match 12.2%; Score 183.5; DB 1; Length 2807;
Best Local Similarity 23.9%; Pred. No. 2.4e-06;
Matches 60; Conservative 19; Mismatches 85; Indels 87; Gaps 7;

QY 2 RGSORVLMMLVLAVGTEHAY-----RPRRVCAVRAHGDVSESFQRYVQPFLLT 54
DB 77 RGOEILRG---PNVCGSRPHSYCCPGWKTLPNGNQICVPICRNSCGDGFCSR---PNMC 130
QY 55 TCDGHRACST--YRTIVRTAYRRSPGLAPRPRVACCPGWKRTSLGFCAGCAAI CQPPCR 112
DB 131 TCSSGQISPTGCRKSIQCCSVRCWNGGT CADDHCQCQKGYIGT-----YCGQPVCEGCGQ 185
QY 113 NGGSCVQPRCPCPAGWRGDTCC----- 135
DB 186 NGRICIGENRCACVYGFTGPQCDYRTGPCFTVNNQMCQGLTGVTKTLCAITGR 245
QY 136 -----SDVDECSARRGGCP-QRCINTAGSYWCQCWE 165
```



```
DR Pfam; PF02210; TSPN; 1.
DR SMART; SM00179; EGF_CA; 3.
DR SMART; SM00282; LamC; 1.
DR SMART; SM00210; TSPN; 1.
DR SMART; SM00214; WVC; 2.
DR PROSITE; PS00010; ASX_HYDROXYL; 3.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; 3.
DR PROSITE; PS00026; EGF_3; 6.
DR PROSITE; PS01187; EGF_CA; 3.
DR PROSITE; PS01208; WVC_1; 2.
DR PROSITE; PS0184; WVC_2; 2.
KW Glycoprotein; EGF-like domain; Repeat; Signal.
FT SIGNAL 1 24
FT CHAIN 25 816
FT DOMAIN 30 258
FT DOMAIN 272 331
FT DOMAIN 332 396
FT DOMAIN 397 439
FT DOMAIN 440 481
FT DOMAIN 482 522
FT DOMAIN 521 553
FT DOMAIN 555 601
FT DOMAIN 602 637
FT DOMAIN 638 693
FT DOMAIN 698 756
FT DOMAIN 758 813
FT DISULFID 401 413
FT DISULFID 407 422
FT DISULFID 424 438
FT DISULFID 444 457
FT DISULFID 451 466
FT DISULFID 468 480
FT DISULFID 486 499
FT DISULFID 493 508
FT DISULFID 510 521
FT DISULFID 525 535
FT DISULFID 529 541
FT DISULFID 543 552
FT DISULFID 559 572
FT DISULFID 566 581
FT DISULFID 583 600
FT DISULFID 606 619
FT DISULFID 613 628
FT DISULFID 630 636
FT CARBOHYD 53 53
FT CARBOHYD 225 225
FT CARBOHYD 293 293
FT CARBOHYD 298 298
FT CARBOHYD 517 517
FT CARBOHYD 615 615
FT CARBOHYD 635 635
SQ SEQUENCE 816 AA; 91163 MW; 5BDD0A94F87E74D CRC64;

Query Match
Best Local Similarity 12.3%; Score 185; DB 1; Length 816;
Matches 47; Conservative 21; Mismatches 64; Indels 28; Gaps 7;

QY 26 PRRVCAVRAHGDPVSEFVORVQPLTT---CDHRACSTYTIYTVARRSPGLAPA 82
DB 461 PGSFMCVKCT-GYRIDDYSCTEHDECLTTHQNCENALC---FNTV-----G 504
QY 83 RPRVACCPGWKRTSLPGACGAAICQPPCRNGSCVQPGRCRCRCPAGWRGDTCCQSDVDECS 142
DB 505 GHNCVCKFGYGN-----GTTCKAFCKDGRNGGACIAANVACRQGGTGPSCETDDECS 560
QY 143 AARGGCPOR--CINTAGSYWCQWEGHSLSDGLICLVKPG 180
DB 561 EGFVQCDSRANCLNLPWGHCECRDGYH---DNGMFAPGG 597

RESULT 4
```



DR SMART; SM00248; ANK; 5.  
DR SMART; SM00179; EGF\_CA; 11.  
DR SMART; SM00004; NL; 3.  
DR PROSITE; PS00297; ANK\_REPEAT; 1.  
DR PROSITE; PS00088; ANK\_REPEAT; 5.  
DR PROSITE; PS00010; ASX\_HYDROXYL; 11.  
DR PROSITE; PS00022; EGF\_1; 28.  
DR PROSITE; PS01186; EGF\_2; 21.  
DR PROSITE; PS00026; EGF\_3; 28.  
DR PROSITE; PS01187; EGF\_CA; 9.  
DR Receptor; transcription regulation; Activator; Differentiation;  
KW Developmental protein; Repeat; ANK repeat; EGF-like domain;  
KW Transmembrane; Glycoprotein; Signal; Phosphorylation; Polymorphism;  
KW Triplet repeat expansion; Alternative splicing.  
FT SIGNAL 1 23 POTENTIAL.  
FT CHAIN 24 2003 NEUROGENIC LOCUS NOTCH PROTEIN HOMOLOG 4.  
FT CHAIN 1432 2003 NOTCH EXTRACELLULAR TRUNCATION  
FT CHAIN 1467 2003 (BY SIMILARITY).  
FT CHAIN 1467 2003 NOTCH INTRACELLULAR DOMAIN  
FT CHAIN 1447 2003 (BY SIMILARITY).  
FT CHAIN 1448 2003 EXTRACELLULAR (POTENTIAL).  
FT CHAIN 1469 2003 POTENTIAL.  
FT CHAIN 24 63 CYTOPLASMIC (POTENTIAL).  
FT CHAIN 64 115 EGF-LIKE 1.  
FT CHAIN 118 155 EGF-LIKE 2.  
FT CHAIN 156 192 EGF-LIKE 3.  
FT CHAIN 194 232 EGF-LIKE 4.  
FT CHAIN 234 274 EGF-LIKE 5.  
FT CHAIN 276 312 EGF-LIKE 6.  
FT CHAIN 314 353 EGF-LIKE 7.  
FT CHAIN 355 391 EGF-LIKE 8.  
FT CHAIN 392 430 EGF-LIKE 9.  
FT CHAIN 432 473 EGF-LIKE 10.  
FT CHAIN 475 511 EGF-LIKE 11.  
FT CHAIN 513 549 EGF-LIKE 12.  
FT CHAIN 551 587 EGF-LIKE 13.  
FT CHAIN 589 625 EGF-LIKE 14.  
FT CHAIN 626 659 EGF-LIKE 15.  
FT CHAIN 661 689 EGF-LIKE 16.  
FT CHAIN 691 727 EGF-LIKE 17.  
FT CHAIN 729 765 EGF-LIKE 18.  
FT CHAIN 767 803 EGF-LIKE 19.  
FT CHAIN 806 842 EGF-LIKE 20.  
FT CHAIN 844 880 EGF-LIKE 21.  
FT CHAIN 882 928 EGF-LIKE 22.  
FT CHAIN 930 966 EGF-LIKE 23.  
FT CHAIN 968 1004 EGF-LIKE 24.  
FT CHAIN 1006 1044 EGF-LIKE 25.  
FT CHAIN 1046 1085 EGF-LIKE 26.  
FT CHAIN 1087 1126 EGF-LIKE 27.  
FT CHAIN 1130 1171 EGF-LIKE 28.  
FT CHAIN 1172 1212 EGF-LIKE 29.  
FT CHAIN 1212 1246 POLY-ARG.  
FT CHAIN 1246 1286 LIN/NOTCH 1.  
FT CHAIN 1286 1365 LIN/NOTCH 2.  
FT CHAIN 1365 1666 ANK 1.  
FT CHAIN 1666 1732 ANK 2.  
FT CHAIN 1732 1765 ANK 3.  
FT CHAIN 1765 1798 ANK 4.  
FT CHAIN 1798 1821 ANK 5.  
FT CHAIN 1821 1846 BY SIMILARITY.  
FT CHAIN 1846 1865 BY SIMILARITY.  
FT CHAIN 1865 1884 BY SIMILARITY.  
FT CHAIN 1884 1903 BY SIMILARITY.  
FT CHAIN 1903 1922 BY SIMILARITY.  
FT CHAIN 1922 1941 BY SIMILARITY.  
FT CHAIN 1941 1960 BY SIMILARITY.  
FT CHAIN 1960 1979 BY SIMILARITY.  
FT CHAIN 1979 1998 BY SIMILARITY.  
FT CHAIN 1998 2017 BY SIMILARITY.  
FT CHAIN 2017 2036 BY SIMILARITY.  
FT CHAIN 2036 2055 BY SIMILARITY.  
FT CHAIN 2055 2074 BY SIMILARITY.  
FT CHAIN 2074 2093 BY SIMILARITY.  
FT CHAIN 2093 2112 BY SIMILARITY.  
FT CHAIN 2112 2131 BY SIMILARITY.  
FT CHAIN 2131 2150 BY SIMILARITY.  
FT CHAIN 2150 2169 BY SIMILARITY.  
FT CHAIN 2169 2188 BY SIMILARITY.  
FT CHAIN 2188 2207 BY SIMILARITY.  
FT CHAIN 2207 2226 BY SIMILARITY.  
FT CHAIN 2226 2245 BY SIMILARITY.  
FT CHAIN 2245 2264 BY SIMILARITY.  
FT CHAIN 2264 2283 BY SIMILARITY.  
FT CHAIN 2283 2302 BY SIMILARITY.  
FT CHAIN 2302 2321 BY SIMILARITY.  
FT CHAIN 2321 2340 BY SIMILARITY.  
FT CHAIN 2340 2359 BY SIMILARITY.  
FT CHAIN 2359 2378 BY SIMILARITY.  
FT CHAIN 2378 2397 BY SIMILARITY.  
FT CHAIN 2397 2416 BY SIMILARITY.  
FT CHAIN 2416 2435 BY SIMILARITY.  
FT CHAIN 2435 2454 BY SIMILARITY.  
FT CHAIN 2454 2473 BY SIMILARITY.  
FT CHAIN 2473 2492 BY SIMILARITY.  
FT CHAIN 2492 2511 BY SIMILARITY.  
FT CHAIN 2511 2530 BY SIMILARITY.  
FT CHAIN 2530 2549 BY SIMILARITY.  
FT CHAIN 2549 2568 BY SIMILARITY.  
FT CHAIN 2568 2587 BY SIMILARITY.  
FT CHAIN 2587 2606 BY SIMILARITY.  
FT CHAIN 2606 2625 BY SIMILARITY.  
FT CHAIN 2625 2644 BY SIMILARITY.  
FT CHAIN 2644 2663 BY SIMILARITY.  
FT CHAIN 2663 2682 BY SIMILARITY.  
FT CHAIN 2682 2701 BY SIMILARITY.  
FT CHAIN 2701 2720 BY SIMILARITY.  
FT CHAIN 2720 2739 BY SIMILARITY.  
FT CHAIN 2739 2758 BY SIMILARITY.  
FT CHAIN 2758 2777 BY SIMILARITY.  
FT CHAIN 2777 2796 BY SIMILARITY.  
FT CHAIN 2796 2815 BY SIMILARITY.  
FT CHAIN 2815 2834 BY SIMILARITY.  
FT CHAIN 2834 2853 BY SIMILARITY.  
FT CHAIN 2853 2872 BY SIMILARITY.  
FT CHAIN 2872 2891 BY SIMILARITY.  
FT CHAIN 2891 2910 BY SIMILARITY.  
FT CHAIN 2910 2929 BY SIMILARITY.  
FT CHAIN 2929 2948 BY SIMILARITY.  
FT CHAIN 2948 2967 BY SIMILARITY.  
FT CHAIN 2967 2986 BY SIMILARITY.  
FT CHAIN 2986 3005 BY SIMILARITY.  
FT CHAIN 3005 3024 BY SIMILARITY.  
FT CHAIN 3024 3043 BY SIMILARITY.  
FT CHAIN 3043 3062 BY SIMILARITY.  
FT CHAIN 3062 3081 BY SIMILARITY.  
FT CHAIN 3081 3100 BY SIMILARITY.  
FT CHAIN 3100 3119 BY SIMILARITY.  
FT CHAIN 3119 3138 BY SIMILARITY.  
FT CHAIN 3138 3157 BY SIMILARITY.  
FT CHAIN 3157 3176 BY SIMILARITY.  
FT CHAIN 3176 3195 BY SIMILARITY.  
FT CHAIN 3195 3214 BY SIMILARITY.  
FT CHAIN 3214 3233 BY SIMILARITY.  
FT CHAIN 3233 3252 BY SIMILARITY.  
FT CHAIN 3252 3271 BY SIMILARITY.  
FT CHAIN 3271 3290 BY SIMILARITY.  
FT CHAIN 3290 3309 BY SIMILARITY.  
FT CHAIN 3309 3328 BY SIMILARITY.  
FT CHAIN 3328 3347 BY SIMILARITY.  
FT CHAIN 3347 3366 BY SIMILARITY.  
FT CHAIN 3366 3385 BY SIMILARITY.  
FT CHAIN 3385 3404 BY SIMILARITY.  
FT CHAIN 3404 3423 BY SIMILARITY.  
FT CHAIN 3423 3442 BY SIMILARITY.  
FT CHAIN 3442 3461 BY SIMILARITY.  
FT CHAIN 3461 3480 BY SIMILARITY.  
FT CHAIN 3480 3500 BY SIMILARITY.  
FT CHAIN 3500 3519 BY SIMILARITY.  
FT CHAIN 3519 3538 BY SIMILARITY.  
FT CHAIN 3538 3557 BY SIMILARITY.  
FT CHAIN 3557 3576 BY SIMILARITY.  
FT CHAIN 3576 3595 BY SIMILARITY.  
FT CHAIN 3595 3614 BY SIMILARITY.  
FT CHAIN 3614 3633 BY SIMILARITY.  
FT CHAIN 3633 3652 BY SIMILARITY.  
FT CHAIN 3652 3671 BY SIMILARITY.  
FT CHAIN 3671 3690 BY SIMILARITY.  
FT CHAIN 3690 3709 BY SIMILARITY.  
FT CHAIN 3709 3728 BY SIMILARITY.  
FT CHAIN 3728 3747 BY SIMILARITY.  
FT CHAIN 3747 3766 BY SIMILARITY.  
FT CHAIN 3766 3785 BY SIMILARITY.  
FT CHAIN 3785 3804 BY SIMILARITY.  
FT CHAIN 3804 3823 BY SIMILARITY.  
FT CHAIN 3823 3842 BY SIMILARITY.  
FT CHAIN 3842 3861 BY SIMILARITY.  
FT CHAIN 3861 3880 BY SIMILARITY.  
FT CHAIN 3880 3900 BY SIMILARITY.  
FT CHAIN 3900 3919 BY SIMILARITY.  
FT CHAIN 3919 3938 BY SIMILARITY.  
FT CHAIN 3938 3957 BY SIMILARITY.  
FT CHAIN 3957 3976 BY SIMILARITY.  
FT CHAIN 3976 3995 BY SIMILARITY.  
FT CHAIN 3995 4014 BY SIMILARITY.  
FT CHAIN 4014 4033 BY SIMILARITY.  
FT CHAIN 4033 4052 BY SIMILARITY.  
FT CHAIN 4052 4071 BY SIMILARITY.  
FT CHAIN 4071 4090 BY SIMILARITY.  
FT CHAIN 4090 4109 BY SIMILARITY.  
FT CHAIN 4109 4128 BY SIMILARITY.  
FT CHAIN 4128 4147 BY SIMILARITY.  
FT CHAIN 4147 4166 BY SIMILARITY.  
FT CHAIN 4166 4185 BY SIMILARITY.  
FT CHAIN 4185 4204 BY SIMILARITY.  
FT CHAIN 4204 4223 BY SIMILARITY.  
FT CHAIN 4223 4242 BY SIMILARITY.  
FT CHAIN 4242 4261 BY SIMILARITY.  
FT CHAIN 4261 4280 BY SIMILARITY.  
FT CHAIN 4280 4300 BY SIMILARITY.  
FT CHAIN 4300 4319 BY SIMILARITY.  
FT CHAIN 4319 4338 BY SIMILARITY.  
FT CHAIN 4338 4357 BY SIMILARITY.  
FT CHAIN 4357 4376 BY SIMILARITY.  
FT CHAIN 4376 4395 BY SIMILARITY.  
FT CHAIN 4395 4414 BY SIMILARITY.  
FT CHAIN 4414 4433 BY SIMILARITY.  
FT CHAIN 4433 4452 BY SIMILARITY.  
FT CHAIN 4452 4471 BY SIMILARITY.  
FT CHAIN 4471 4490 BY SIMILARITY.  
FT CHAIN 4490 4509 BY SIMILARITY.  
FT CHAIN 4509 4528 BY SIMILARITY.  
FT CHAIN 4528 4547 BY SIMILARITY.  
FT CHAIN 4547 4566 BY SIMILARITY.  
FT CHAIN 4566 4585 BY SIMILARITY.  
FT CHAIN 4585 4604 BY SIMILARITY.  
FT CHAIN 4604 4623 BY SIMILARITY.  
FT CHAIN 4623 4642 BY SIMILARITY.  
FT CHAIN 4642 4661 BY SIMILARITY.  
FT CHAIN 4661 4680 BY SIMILARITY.  
FT CHAIN 4680 4700 BY SIMILARITY.  
FT CHAIN 4700 4719 BY SIMILARITY.  
FT CHAIN 4719 4738 BY SIMILARITY.  
FT CHAIN 4738 4757 BY SIMILARITY.  
FT CHAIN 4757 4776 BY SIMILARITY.  
FT CHAIN 4776 4795 BY SIMILARITY.  
FT CHAIN 4795 4814 BY SIMILARITY.  
FT CHAIN 4814 4833 BY SIMILARITY.  
FT CHAIN 4833 4852 BY SIMILARITY.  
FT CHAIN 4852 4871 BY SIMILARITY.  
FT CHAIN 4871 4890 BY SIMILARITY.  
FT CHAIN 4890 4909 BY SIMILARITY.  
FT CHAIN 4909 4928 BY SIMILARITY.  
FT CHAIN 4928 4947 BY SIMILARITY.  
FT CHAIN 4947 4966 BY SIMILARITY.  
FT CHAIN 4966 4985 BY SIMILARITY.  
FT CHAIN 4985 5004 BY SIMILARITY.  
FT CHAIN 5004 5023 BY SIMILARITY.  
FT CHAIN 5023 5042 BY SIMILARITY.  
FT CHAIN 5042 5061 BY SIMILARITY.  
FT CHAIN 5061 5080 BY SIMILARITY.  
FT CHAIN 5080 5100 BY SIMILARITY.  
FT CHAIN 5100 5119 BY SIMILARITY.  
FT CHAIN 5119 5138 BY SIMILARITY.  
FT CHAIN 5138 5157 BY SIMILARITY.  
FT CHAIN 5157 5176 BY SIMILARITY.  
FT CHAIN 5176 5195 BY SIMILARITY.  
FT CHAIN 5195 5214 BY SIMILARITY.  
FT CHAIN 5214 5233 BY SIMILARITY.  
FT CHAIN 5233 5252 BY SIMILARITY.  
FT CHAIN 5252 5271 BY SIMILARITY.  
FT CHAIN 5271 5290 BY SIMILARITY.  
FT CHAIN 5290 5309 BY SIMILARITY.  
FT CHAIN 5309 5328 BY SIMILARITY.  
FT CHAIN 5328 5347 BY SIMILARITY.  
FT CHAIN 5347 5366 BY SIMILARITY.  
FT CHAIN 5366 5385 BY SIMILARITY.  
FT CHAIN 5385 5404 BY SIMILARITY.  
FT CHAIN 5404 5423 BY SIMILARITY.  
FT CHAIN 5423 5442 BY SIMILARITY.  
FT CHAIN 5442 5461 BY SIMILARITY.  
FT CHAIN 5461 5480 BY SIMILARITY.  
FT CHAIN 5480 5500 BY SIMILARITY.  
FT CHAIN 5500 5519 BY SIMILARITY.  
FT CHAIN 5519 5538 BY SIMILARITY.  
FT CHAIN 5538 5557 BY SIMILARITY.  
FT CHAIN 5557 5576 BY SIMILARITY.  
FT CHAIN 5576 5595 BY SIMILARITY.  
FT CHAIN 5595 5614 BY SIMILARITY.  
FT CHAIN 5614 5633 BY SIMILARITY.  
FT CHAIN 5633 5652 BY SIMILARITY.  
FT CHAIN 5652 5671 BY SIMILARITY.  
FT CHAIN 5671 5690 BY SIMILARITY.  
FT CHAIN 5690 5709 BY SIMILARITY.  
FT CHAIN 5709 5728 BY SIMILARITY.  
FT CHAIN 5728 5747 BY SIMILARITY.  
FT CHAIN 5747 5766 BY SIMILARITY.  
FT CHAIN 5766 5785 BY SIMILARITY.  
FT CHAIN 5785 5804 BY SIMILARITY.  
FT CHAIN 5804 5823 BY SIMILARITY.  
FT CHAIN 5823 5842 BY SIMILARITY.  
FT CHAIN 5842 5861 BY SIMILARITY.  
FT CHAIN 5861 5880 BY SIMILARITY.  
FT CHAIN 5880 5900 BY SIMILARITY.  
FT CHAIN 5900 5919 BY SIMILARITY.  
FT CHAIN 5919 5938 BY SIMILARITY.  
FT CHAIN 5938 5957 BY SIMILARITY.  
FT CHAIN 5957 5976 BY SIMILARITY.  
FT CHAIN 5976 5995 BY SIMILARITY.  
FT CHAIN 5995 6014 BY SIMILARITY.  
FT CHAIN 6014 6033 BY SIMILARITY.  
FT CHAIN 6033 6052 BY SIMILARITY.  
FT CHAIN 6052 6071 BY SIMILARITY.  
FT CHAIN 6071 6090 BY SIMILARITY.  
FT CHAIN 6090 6109 BY SIMILARITY.  
FT CHAIN 6109 6128 BY SIMILARITY.  
FT CHAIN 6128 6147 BY SIMILARITY.  
FT CHAIN 6147 6166 BY SIMILARITY.  
FT CHAIN 6166 6185 BY SIMILARITY.  
FT CHAIN 6185 6204 BY SIMILARITY.  
FT CHAIN 6204 6223 BY SIMILARITY.  
FT CHAIN 6223 6242 BY SIMILARITY.  
FT CHAIN 6242 6261 BY SIMILARITY.  
FT CHAIN 6261 6280 BY SIMILARITY.  
FT CHAIN 6280 6300 BY SIMILARITY.  
FT CHAIN 6300 6319 BY SIMILARITY.  
FT CHAIN 6319 6338 BY SIMILARITY.  
FT CHAIN 6338 6357 BY SIMILARITY.  
FT CHAIN 6357 6376 BY SIMILARITY.  
FT CHAIN 6376 6395 BY SIMILARITY.  
FT CHAIN 6395 6414 BY SIMILARITY.  
FT CHAIN 6414 6433 BY SIMILARITY.  
FT CHAIN 6433 6452 BY SIMILARITY.  
FT CHAIN 6452 6471 BY SIMILARITY.  
FT CHAIN 6471 6490 BY SIMILARITY.  
FT CHAIN 6490 6509 BY SIMILARITY.  
FT CHAIN 6509 6528 BY SIMILARITY.  
FT CHAIN 6528 6547 BY SIMILARITY.  
FT CHAIN 6547 6566 BY SIMILARITY.  
FT CHAIN 6566 6585 BY SIMILARITY.  
FT CHAIN 6585 6604 BY SIMILARITY.  
FT CHAIN 6604 6623 BY SIMILARITY.  
FT CHAIN 6623 6642 BY SIMILARITY.  
FT CHAIN 6642 6661 BY SIMILARITY.  
FT CHAIN 6661 6680 BY SIMILARITY.  
FT CHAIN 6680 6700 BY SIMILARITY.  
FT CHAIN 6700 6719 BY SIMILARITY.  
FT CHAIN 6719 6738 BY SIMILARITY.  
FT CHAIN 6738 6757 BY SIMILARITY.  
FT CHAIN 6757 6776 BY SIMILARITY.  
FT CHAIN 6776 6795 BY SIMILARITY.  
FT CHAIN 6795 6814 BY SIMILARITY.  
FT CHAIN 6814 6833 BY SIMILARITY.  
FT CHAIN 6833 6852 BY SIMILARITY.  
FT CHAIN 6852 6871 BY SIMILARITY.  
FT CHAIN 6871 6890 BY SIMILARITY.  
FT CHAIN 6890 6909 BY SIMILARITY.  
FT CHAIN 6909 6928 BY SIMILARITY.  
FT CHAIN 6928 6947 BY SIMILARITY.  
FT CHAIN 6947 6966 BY SIMILARITY.  
FT CHAIN 6966 6985 BY SIMILARITY.  
FT CHAIN 6985 7004 BY SIMILARITY.  
FT CHAIN 7004 7023 BY SIMILARITY.  
FT CHAIN 7023 7042 BY SIMILARITY.  
FT CHAIN 7042 7061 BY SIMILARITY.  
FT CHAIN 7061 7080 BY SIMILARITY.  
FT CHAIN 7080 7100 BY SIMILARITY.  
FT CHAIN 7100 7119 BY SIMILARITY.  
FT CHAIN 7119 7138 BY SIMILARITY.  
FT CHAIN 7138 7157 BY SIMILARITY.  
FT CHAIN 7157 7176 BY SIMILARITY.  
FT CHAIN 7176 7195 BY SIMILARITY.  
FT CHAIN 7195 7214 BY SIMILARITY.  
FT CHAIN 7214 7233 BY SIMILARITY.  
FT CHAIN 7233 7252 BY SIMILARITY.  
FT CHAIN 7252 7271 BY SIMILARITY.  
FT CHAIN 7271 7290 BY SIMILARITY.  
FT CHAIN 7290 7309 BY SIMILARITY.  
FT CHAIN 7309 7328 BY SIMILARITY.  
FT CHAIN 7328 7347 BY SIMILARITY.  
FT CHAIN 7347 7366 BY SIMILARITY.  
FT CHAIN 7366 7385 BY SIMILARITY.  
FT CHAIN 7385 7404 BY SIMILARITY.  
FT CHAIN 7404 7423 BY SIMILARITY.  
FT CHAIN 7423 7442 BY SIMILARITY.  
FT CHAIN 7442 7461 BY SIMILARITY.  
FT CHAIN 7461 7480 BY SIMILARITY.  
FT CHAIN 7480 7500 BY SIMILARITY.  
FT CHAIN 7500 7519 BY SIMILARITY.  
FT CHAIN 7519 7538 BY SIMILARITY.  
FT CHAIN 7538 7557 BY SIMILARITY.  
FT CHAIN 7557 7576 BY SIMILARITY.  
FT CHAIN 7576 7595 BY SIMILARITY.  
FT CHAIN 7595 7614 BY SIMILARITY.  
FT CHAIN 7614 7633 BY SIMILARITY.  
FT CHAIN 7633 7652 BY SIMILARITY.  
FT CHAIN 7652 7671 BY SIMILARITY.  
FT CHAIN 7671 7690 BY SIMILARITY.  
FT CHAIN 7690 7709 BY SIMILARITY.  
FT CHAIN 7709 7728 BY SIMILARITY.  
FT CHAIN 7728 7747 BY SIMILARITY.  
FT CHAIN 7747 7766 BY SIMILARITY.  
FT CHAIN 7766 7785 BY SIMILARITY.  
FT CHAIN 7785 7804 BY SIMILARITY.  
FT CHAIN 7804 7823 BY SIMILARITY.  
FT CHAIN 7823 7842 BY SIMILARITY.  
FT CHAIN 7842 7861 BY SIMILARITY.  
FT CHAIN 7861 7880 BY SIMILARITY.  
FT CHAIN 7880 7900 BY SIMILARITY.  
FT CHAIN 7900 7919 BY SIMILARITY.  
FT CHAIN 7919 7938 BY SIMILARITY.  
FT CHAIN 7938 7957 BY SIMILARITY.  
FT CHAIN 7957 7976 BY SIMILARITY.  
FT CHAIN 7976 7995 BY SIMILARITY.  
FT CHAIN 7995 8014 BY SIMILARITY.  
FT CHAIN 8014 8033 BY SIMILARITY.  
FT CHAIN 8033 8052 BY SIMILARITY.  
FT CHAIN 8052 8071 BY SIMILARITY.  
FT CHAIN 8071 8090 BY SIMILARITY.  
FT CHAIN 8090 8109 BY SIMILARITY.  
FT CHAIN 8109 8128 BY SIMILARITY.  
FT CHAIN 8128 8147 BY SIMILARITY.  
FT CHAIN 8147 8166 BY SIMILARITY.  
FT CHAIN 8166 8185 BY SIMILARITY.  
FT CHAIN 8185 8204 BY SIMILARITY.  
FT CHAIN 8204 8223 BY SIMILARITY.  
FT CHAIN 8223 8242 BY SIMILARITY.  
FT CHAIN 8242 8261 BY SIMILARITY.  
FT CHAIN 8261 8280 BY SIMILARITY.  
FT CHAIN 8280 8300 BY SIMILARITY.  
FT CHAIN 8300 8319 BY SIMILARITY.  
FT CHAIN 8319 8338 BY SIMILARITY.  
FT CHAIN 8338 8357 BY SIMILARITY.  
FT CHAIN 8357 8376 BY SIMILARITY.  
FT CHAIN 8376 8395 BY SIMILARITY.  
FT CHAIN 8395 8414 BY SIMILARITY.  
FT CHAIN 8414 8433 BY SIMILARITY.  
FT CHAIN 8433 8452 BY SIMILARITY.  
FT CHAIN 8452 8471 BY SIMILARITY.  
FT CHAIN 8471 8490 BY SIMILARITY.  
FT CHAIN 8490 8509 BY SIMILARITY.  
FT CHAIN 8509 8528 BY SIMILARITY.  
FT CHAIN 8528 8547 BY SIMILARITY.  
FT CHAIN 8547 8566 BY SIMILARITY.  
FT CHAIN 8566 8585 BY SIMILARITY.  
FT CHAIN 8585 8604 BY SIMILARITY.  
FT CHAIN 8604 8623 BY SIMILARITY.  
FT CHAIN 8623 8642 BY SIMILARITY.  
FT CHAIN 8642 8661 BY SIMILARITY.  
FT CHAIN 8661 8680 BY SIMILARITY.  
FT CHAIN 8680 8700 BY SIMILARITY.  
FT CHAIN 8700 8719 BY SIMILARITY.  
FT CHAIN 8719 8738 BY SIMILARITY.  
FT CHAIN 8738 8757 BY SIMILARITY.  
FT CHAIN 8757 8776 BY SIMILARITY.  
FT CHAIN 8776 8795 BY SIMILARITY.  
FT CHAIN 8795 8814 BY SIMILARITY.  
FT CHAIN 8814 8833 BY SIMILARITY.  
FT CHAIN 8833 8852 BY SIMILARITY.  
FT CHAIN 8852 8871 BY SIMILARITY.  
FT CHAIN 8871 8890 BY SIMILARITY.  
FT CHAIN 8890 8909 BY SIMILARITY.  
FT CHAIN 8909 8928 BY SIMILARITY.  
FT CHAIN 8928 8947 BY SIMILARITY.  
FT CHAIN 8947 8966 BY SIMILARITY.  
FT CHAIN 8966 8985 BY SIMILARITY.  
FT CHAIN 8985 9004 BY SIMILARITY.  
FT CHAIN 9004 9023 BY SIMILARITY.  
FT CHAIN 9023 9042 BY SIMILARITY.  
FT CHAIN 9042 9061 BY SIMILARITY.  
FT CHAIN 9061 9080 BY SIMILARITY.  
FT CHAIN 9080 9100 BY SIMILARITY.  
FT CHAIN 9100 9119 BY SIMILARITY.  
FT CHAIN 9119 9138 BY SIMILARITY.  
FT CHAIN 9138 9157 BY SIMILARITY.  
FT CHAIN 9157 9176 BY SIMILARITY.  
FT CHAIN 9176 9195 BY SIMILARITY.  
FT CHAIN 9195 9214 BY SIMILARITY.  
FT CHAIN 9214 9233 BY SIMILARITY.  
FT CHAIN 9233 9252 BY SIMILARITY.  
FT CHAIN 9252 9271 BY SIMILARITY.  
FT CHAIN 9271 9290 BY SIMILARITY.  
FT CHAIN 9290 9309 BY SIMILARITY.  
FT CHAIN 9309 9328 BY SIMILARITY.  
FT CHAIN 9328 9347 BY SIMILARITY.  
FT CHAIN 9347 9366 BY SIMILARITY.  
FT CHAIN 9366 9385 BY SIMILARITY.  
FT CHAIN 9385 9404 BY SIMILARITY.  
FT CHAIN 9404 9423 BY SIMILARITY.  
FT CHAIN 9423 9442 BY SIMILARITY.  
FT CHAIN 9442 9461 BY SIMILARITY.  
FT CHAIN 9461 9480 BY SIMILARITY.  
FT CHAIN 9480 9500 BY SIMILARITY.  
FT CHAIN 9500 9519 BY SIMILARITY.  
FT CHAIN 9519 9538 BY SIMILARITY.  
FT CHAIN 9538 9557 BY SIMILARITY.  
FT CHAIN 9557 9576 BY SIMILARITY.  
FT CHAIN 9576 9595 BY SIMILARITY.  
FT CHAIN 9595 9614 BY SIMILARITY.  
FT CHAIN 9614 9633 BY SIMILARITY.  
FT CHAIN 9633 9652 BY SIMILARITY.  
FT CHAIN 9652 9671 BY SIMILARITY.  
FT CHAIN 9671 9690 BY SIMILARITY.  
FT CHAIN 9690 9709 BY SIMILARITY.  
FT CHAIN 9709 9728 BY SIMILARITY.  
FT CHAIN 9728 9747 BY SIMILARITY.  
FT CHAIN 9747 9766 BY SIMILARITY.  
FT CHAIN 9766 9785 BY SIMILARITY.  
FT CHAIN 9785 9804 BY SIMILARITY.  
FT CHAIN 9804 9823 BY SIMILARITY.  
FT CHAIN 9823 9842 BY SIMILARITY.  
FT CHAIN 9842 9861 BY SIMILARITY.  
FT CHAIN 9861 9880 BY SIMILARITY.  
FT CHAIN 9880 9900 BY SIMILARITY.  
FT CHAIN 9900 9919 BY SIMILARITY.  
FT CHAIN 9919 9938 BY SIMILARITY.  
FT CHAIN 9938 9957 BY SIMILARITY.  
FT CHAIN 9957 9976 BY SIMILARITY.  
FT CHAIN 9976 9995 BY SIMILARITY.  
FT CHAIN 9995 10014 BY SIMILARITY.  
FT CHAIN 10014 10033 BY SIMILARITY.  
FT CHAIN 10033 10052 BY SIMILARITY.  
FT CHAIN 10052 10071 BY SIMILARITY.  
FT CHAIN 10071 10090 BY SIMILARITY.  
FT CHAIN 10090 10109 BY SIMILARITY.  
FT CHAIN 10109 10128 BY SIMILARITY.  
FT CHAIN 10128 10147 BY SIMILARITY.  
FT CHAIN 10147 10166 BY SIMILARITY.  
FT CHAIN 10166 10185 BY SIMILARITY.  
FT CHAIN 10185 10204 BY SIMILARITY.  
FT CHAIN 10204 10223 BY SIMILARITY.  
FT CHAIN 10223 10242 BY SIMILARITY.  
FT CHAIN 10242 10261 BY SIMILARITY.  
FT CHAIN 10261 10280 BY SIMILARITY.  
FT CHAIN 10280 10300 BY SIMILARITY.  
FT CHAIN 10300 10319 BY SIMILARITY.  
FT CHAIN 10319 10338 BY SIMILARITY.  
FT CHAIN 10338 10357 BY SIMILARITY.  
FT CHAIN 10357 10376 BY SIMILARITY.  
FT CHAIN 10376 10395 BY SIMILARITY.  
FT CHAIN 10395 10414 BY SIMILARITY.  
FT CHAIN 10414 10433 BY SIMILARITY.  
FT CHAIN 10433 10452 BY SIMILARITY.  
FT CHAIN 10452 10471 BY SIMILARITY.  
FT CHAIN 10471 10490 BY SIMILARITY.  
FT CHAIN 10490 10509 BY SIMILARITY.  
FT CHAIN 10509 10528 BY SIMILARITY.  
FT CHAIN 10528 10547 BY SIMILARITY.  
FT CHAIN 10547 10566 BY SIMILARITY.  
FT CHAIN 10566 10585 BY SIMILARITY.  
FT CHAIN 10585 10604 BY SIMILARITY.  
FT CHAIN 10604 10623 BY SIMILARITY.  
FT CHAIN 10623 10642 BY SIMILARITY.  
FT CHAIN 10642 10661 BY SIMILARITY.  
FT CHAIN 10661 10680 BY SIMILARITY.  
FT CHAIN 10680 10700 BY SIMILARITY.  
FT CHAIN 10700 10719 BY SIMILARITY.  
FT CHAIN 10719 10738 BY SIMILARITY.  
FT CHAIN 10738 10757 BY SIMILARITY.  
FT CHAIN 10757 10776 BY SIMILARITY.  
FT CHAIN 10776 10795 BY SIMILARITY.  
FT CHAIN 10795 10814 BY SIMILARITY.  
FT CHAIN 10814 10833 BY SIMILARITY.  
FT CHAIN 10833 10852 BY SIMILARITY.  
FT CHAIN 10852 10871 BY SIMILARITY.  
FT CHAIN 10871 10890 BY SIMILARITY.  
FT CHAIN 10890 10909 BY SIMILARITY.  
FT CHAIN 10909 10928 BY SIMILARITY.  
FT CHAIN 10928 10947 BY SIMILARITY.  
FT CHAIN 10947 10966 BY SIMILARITY.  
FT CHAIN 10966 10985 BY SIMILARITY.  
FT CHAIN 10985 11004 BY SIMILARITY.  
FT CHAIN 11004 11023 BY SIMILARITY.  
FT CHAIN 11023 11042 BY SIMILARITY.  
FT CHAIN 11042 11061 BY SIMILARITY.  
FT CHAIN 11061 11080 BY SIMILARITY.  
FT CHAIN 11080 11100 BY SIMILARITY.  
FT CHAIN 11100 11119 BY SIMILARITY.  
FT CHAIN 11119 11138 BY SIMILARITY.  
FT CHAIN 11138 11157 BY SIMILARITY.  
FT CHAIN 11157 11176 BY SIMILARITY.  
FT CHAIN 11176 11195 BY SIMILARITY.  
FT CHAIN 11195 11214 BY SIMILARITY.  
FT CHAIN 11214 11233 BY SIMILARITY.  
FT CHAIN 11233 11252 BY SIMILARITY.  
FT CHAIN 11252 11271 BY SIMILARITY.  
FT CHAIN 11271 11290 BY SIMILARITY.  
FT CHAIN 11290 11309 BY SIMILARITY.  
FT CHAIN 11309 11328 BY SIMILARITY.  
FT CHAIN 11328 11347 BY SIMILARITY.  
FT CHAIN 11347 11366 BY SIMILARITY.  
FT CHAIN 11366 11385 BY SIMILARITY.  
FT CHAIN 11385 11404 BY SIMILARITY.  
FT CHAIN 11404 11423 BY SIMILARITY.  
FT CHAIN 11423 11442 BY SIMILARITY.  
FT CHAIN 11442 11461 BY SIMILARITY.  
FT CHAIN 11461 11480 BY SIMILARITY.  
FT CHAIN 11480 11500 BY SIMILARITY.  
FT CHAIN 11500 11519 BY SIMILARITY.  
FT CHAIN 11519 11538 BY SIMILARITY.  
FT CHAIN 11538 11557 BY SIMILARITY.  
FT CHAIN 11557 11576 BY SIMILARITY.  
FT CHAIN 11576 11595 BY SIMILARITY.  
FT CHAIN 11595 11614 BY SIMILARITY.  
FT CHAIN 11614 11633 BY SIMILARITY.<



RT Notch4 in embryonic endothelium.";

RL Proc. Natl. Acad. Sci. U.S.A. 98:5643-5648 (2001).

RN [7]

RP SEQUENCE OF 1463-1964, POST-TRANSLATIONAL PROCESSING, AND MUTAGENESIS

RP OF VAL-1463.

RX MEDLINE-21523956; PubMed-11518718;

RA Saxena M.T., Schroeter E.H., Mumm J.S., Kopan R.;

RT "Murine notch homologs (NL-4) undergo presenilin-dependent

RT proteolysis.";

RL J. Biol. Chem. 276:40268-40273 (2001).

RN [8]

RP POST-TRANSLATIONAL PROCESSING.

RX MEDLINE-21374376; PubMed-1145941;

RA Mizutani T., Taniuchi Y., Aoki T., Hashimoto N., Honjo T.;

RT "Conservation of the biochemical mechanisms of signal transduction

RT among mammalian Notch family members.";

RL Proc. Natl. Acad. Sci. U.S.A. 98:9026-9031 (2001).

CC -!- FUNCTION: Functions as a receptor for membrane-bound ligands

CC Jagged1, Jagged2 and Delta to regulate cell-fate determination.

CC Upon ligand activation through the released notch intracellular

CC domain (NICD) it forms a transcriptional activator complex with

CC RBP-J kappa and activates genes of the enhancer of split locus.

CC Affects the implementation of differentiation, proliferation and

CC apoptotic programs (by similarity). May regulate branching

CC morphogenesis in the developing vascular system.

CC -!- SUBUNIT: Heterodimer of a C-terminal fragment N(TW) and a N-

CC terminal fragment N(BC) which are probably linked by disulfide

CC bonds.

CC -!- SUBCELLULAR LOCATION: Type I membrane protein. Following

CC proteolytical processing NICD is translocated to the nucleus.

CC -!- TISSUE SPECIFICITY: Highly expressed in lung, moderately in heart

CC kidney, and at lower levels in the ovary and skeletal muscle. A

CC very low expression is seen in the brain, intestine, liver and

CC testis.

CC -!- DEVELOPMENTAL STAGE: Highly expressed in endothelial cells during

CC embryonic development from 9.0 dpc.

CC -!- PTM: Synthesized in the endoplasmic reticulum as an inactive form

CC which is proteolytically cleaved by a furin-like convertase in the

CC trans-Golgi network before it reaches the plasma membrane to yield

CC an active, ligand-accessible form. Cleavage results in a C-

CC terminal fragment N(TW) and a N-terminal fragment N(BC). Following

CC ligand binding, it is cleaved by TNF-alpha converting enzyme

CC (TACE) to yield a membrane-associated intermediate fragment called

CC notch extracellular truncation (NEXT). This fragment is then

CC cleaved by presenilin dependent gamma-secretase to release a

CC notch-derived peptide containing the intracellular domain (NICD)

CC from the membrane.

CC -!- PTM: Phosphorylated.

CC -!- DISEASE: Loss of the extracellular domain causes constitutive

CC activation of the Notch protein, which leads to hyperproliferation

CC of glandular epithelial tissues and development of mammary

CC carcinomas.

CC -!- SIMILARITY: Belongs to the NOTCH family.

CC -!- SIMILARITY: Contains 29 EGF-like domains.

CC -!- SIMILARITY: Contains 3 Lin/Notch repeats.

CC -!- SIMILARITY: Contains 5 ANK repeats.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL; M60456; AAB38377.1; -;

DR EMBL; U43691; AAC52630.1; -;

DR EMBL; U43691; AAC52631.1; -;

DR EMBL; AF030001; AAB82004.1; -;

DR EMBL; AB016771; BAA32281.1; ALT SEQ.

DR EMBL; AB016772; BAA32283.1; ALT\_INIT.

DR EMBL; AB016773; BAA32284.1; ALT\_INIT.

DR EMBL; AB016774; BAA32285.1; -;

DR PIR; A38072; TVNVT3.

DR PIR; T09059; T09059.

DR HSP; P08709; IBF9.

DR MGD; MGI.1107471; Notch4.

DR InterPro; IPR002110; ANK.

DR InterPro; IPR000152; Asx\_hydroxyl\_1\_S.

DR InterPro; IPR000742; EGF\_2.

DR InterPro; IPR001881; EGF\_Ca.

DR InterPro; IPR001438; EGF\_II.

DR InterPro; IPR006209; EGF\_like.

DR InterPro; IPR002049; Laminin\_EGF.

DR InterPro; IPR008297; Notch\_dom.

DR InterPro; IPR000800; Notch\_dom.

DR Pfam; PF00023; ank; 6.

DR Pfam; PF00008; EGF; 27.

DR Pfam; PF00056; notch; 2.

DR PIRSF; PIRSF002279; Notch; 1.

DR PRINTS; PR00010; EGF\_BLOOD.

DR PRINTS; PR00011; EGF\_LAMININ.

DR PRINTS; PR01452; NOTCH.

DR SMART; SM00248; ANK; 6.

DR SMART; SM00179; EGF\_CA; 11.

DR SMART; SM00004; NL; 2.

DR PROSITE; PSS0297; ANK\_REPEAT; 1.

DR PROSITE; PSS0088; ANK\_REPEAT; 5.

DR PROSITE; PSS0010; ASX\_HYDROXYL; 11.

DR PROSITE; PSS0032; EGF\_1; 28.

DR PROSITE; PSS0186; EGF\_2; 21.

DR PROSITE; PSS0026; EGF\_3; 27.

DR PROSITE; PSS0187; EGF\_CA; 9.

DR REceptor; Transcription regulation; Activator; Differentiation;

KW Developmental protein; Repeat; ANK repeat; EGF-like domain;

KW Transmembrane; Glycoprotein; Signal; Phosphorylation; Proto-oncogene.

FT SIGNAL 1 20

FT CHAIN 21 1964

FT CHAIN 1411 1964

FT CHAIN 1428 1964

FT CHAIN 1463 1964

FT DOMAIN 21 1443

FT TRANSMEM 1444 1464

FT DOMAIN 1465 1964

FT DOMAIN 21 60

FT DOMAIN 61 112

FT DOMAIN 115 152

FT DOMAIN 153 189

FT DOMAIN 191 229

FT DOMAIN 231 271

FT DOMAIN 273 309

FT DOMAIN 311 350

FT DOMAIN 352 388

FT DOMAIN 389 427

FT DOMAIN 429 470

FT DOMAIN 472 508

FT DOMAIN 510 546

FT DOMAIN 548 584

FT DOMAIN 586 622

FT DOMAIN 623 656

FT DOMAIN 658 686

FT DOMAIN 688 724

FT DOMAIN 726 762

FT DOMAIN 764 800

FT DOMAIN 803 839

FT DOMAIN 841 877

FT DOMAIN 878 924

FT DOMAIN 926 962

FT DOMAIN 964 1000

FT DOMAIN 1002 1040

FT DOMAIN 1042 1081

FT DOMAIN 1083 1122

FT DOMAIN 1126 1167

FT REPEAT 1168 1208

FT REPEAT 1209 1242

FT REPEAT 1243 1282

FT SIGNAL 1 20

FT CHAIN 21 1964

FT CHAIN 1411 1964

FT CHAIN 1428 1964

FT CHAIN 1463 1964

FT DOMAIN 21 1443

FT TRANSMEM 1444 1464

FT DOMAIN 1465 1964

FT DOMAIN 21 60

FT DOMAIN 61 112

FT DOMAIN 115 152

FT DOMAIN 153 189

FT DOMAIN 191 229

FT DOMAIN 231 271

FT DOMAIN 273 309

FT DOMAIN 311 350

FT DOMAIN 352 388

FT DOMAIN 389 427

FT DOMAIN 429 470

FT DOMAIN 472 508

FT DOMAIN 510 546

FT DOMAIN 548 584

FT DOMAIN 586 622

FT DOMAIN 623 656

FT DOMAIN 658 686

FT DOMAIN 688 724

FT DOMAIN 726 762

FT DOMAIN 764 800

FT DOMAIN 803 839

FT DOMAIN 841 877

FT DOMAIN 878 924

FT DOMAIN 926 962

FT DOMAIN 964 1000

FT DOMAIN 1002 1040

FT DOMAIN 1042 1081

FT DOMAIN 1083 1122

FT DOMAIN 1126 1167

FT REPEAT 1168 1208

FT REPEAT 1209 1242

FT REPEAT 1243 1282

FT SIGNAL 1 20

FT CHAIN 21 1964

FT CHAIN 1411 1964

FT CHAIN 1428 1964

FT CHAIN 1463 1964

FT DOMAIN 21 1443

FT TRANSMEM 1444 1464

FT DOMAIN 1465 1964

FT DOMAIN 21 60

FT DOMAIN 61 112

FT DOMAIN 115 152

FT DOMAIN 153 189

FT DOMAIN 191 229

FT DOMAIN 231 271

FT DOMAIN 273 309

FT DOMAIN 311 350

FT DOMAIN 352 388

FT DOMAIN 389 427

FT DOMAIN 429 470

FT DOMAIN 472 508

FT DOMAIN 510 546

FT DOMAIN 548 584

FT DOMAIN 586 622

FT DOMAIN 623 656

FT DOMAIN 658 686

FT DOMAIN 688 724

FT DOMAIN 726 762

FT DOMAIN 764 800

FT DOMAIN 803 839

FT DOMAIN 841 877

FT DOMAIN 878 924

FT DOMAIN 926 962

FT DOMAIN 964 1000

FT DOMAIN 1002 1040

FT DOMAIN 1042 1081

FT DOMAIN 1083 1122

FT DOMAIN 1126 1167

FT REPEAT 1168 1208

FT REPEAT 1209 1242

FT REPEAT 1243 1282

FT SIGNAL 1 20

FT CHAIN 21 1964

FT CHAIN 1411 1964

FT CHAIN 1428 1964

FT CHAIN 1463 1964

FT DOMAIN 21 1443

FT TRANSMEM 1444 1464

FT DOMAIN 1465 1964

FT DOMAIN 21 60

FT DOMAIN 61 112

FT DOMAIN 115 152

FT DOMAIN 153 189

FT DOMAIN 191 229

FT DOMAIN 231 271

FT DOMAIN 273 309

FT DOMAIN 311 350

FT DOMAIN 352 388

FT DOMAIN 389 427

FT DOMAIN 429 470

FT DOMAIN 472 508

FT DOMAIN 510 546

FT DOMAIN 548 584

FT DOMAIN 586 622

FT DOMAIN 623 656

FT DOMAIN 658 686

FT DOMAIN 688 724

FT DOMAIN 726 762

FT DOMAIN 764 800

FT DOMAIN 803 839

FT DOMAIN 841 877

FT DOMAIN 878 924

FT DOMAIN 926 962

FT DOMAIN 964 1000

FT DOMAIN 1002 1040

FT DOMAIN 1042 1081

FT DOMAIN 1083 1122

FT DOMAIN 1126 1167

FT REPEAT 1168 1208

FT REPEAT 1209 1242

FT REPEAT 1243 1282

FT SIGNAL 1 20

FT CHAIN 21 1964

FT CHAIN 1411 1964

FT CHAIN 1428 1964

FT CHAIN 1463 1964

FT DOMAIN 21 1443

FT TRANSMEM 1444 1464

FT DOMAIN 1465 1964

FT DOMAIN 21 60

FT DOMAIN 61 112

FT DOMAIN 115 152

FT DOMAIN 153 189

FT DOMAIN 191 229

FT DOMAIN 231 271

FT DOMAIN 273 309

FT DOMAIN 311 350

FT DOMAIN 352 388

FT DOMAIN 389 427

FT DOMAIN 429 470

FT DOMAIN 472 508

FT DOMAIN 510 546

FT DOMAIN 548 584

FT DOMAIN 586 622

FT DOMAIN 623 656

FT DOMAIN 658 686

FT DOMAIN 688 724

FT DOMAIN 726 762

FT DOMAIN 764 800

FT DOMAIN 803 839

FT DOMAIN 841 877

FT DOMAIN 878 924

FT DOMAIN 926 962

FT DOMAIN 964 1000

FT DOMAIN 1002 1040

FT DOMAIN 1042 1081

FT DOMAIN 1083 1122

FT DOMAIN 1126 1167

FT REPEAT 1168 1208

FT REPEAT 1209 1242

FT REPEAT 1243 1282

FT SIGNAL 1 20

FT CHAIN 21 1964

FT CHAIN 1411 1964

FT CHAIN 1428 1964

FT CHAIN 1463 1964

FT DOMAIN 21 1443

FT TRANSMEM 1444 1464

FT DOMAIN 1465 1964

FT DOMAIN 21 60

FT DOMAIN 61 112

FT DOMAIN 115 152

FT DOMAIN 153 189

FT DOMAIN 191 229

FT DOMAIN 231 271

FT DOMAIN 273 309

FT DOMAIN 311 350

FT DOMAIN 352 388

FT DOMAIN 389 427

FT DOMAIN 429 470

FT DOMAIN 472 508

FT DOMAIN 510 546

FT DOMAIN 548 584

FT DOMAIN 586 622

FT DOMAIN 623 656

FT DOMAIN 658 686

FT DOMAIN 688 724

FT DOMAIN 726 762

FT DOMAIN 764 800

FT DOMAIN 803 839

FT DOMAIN 841 877

FT DOMAIN 878 924

FT DOMAIN 926 962

FT DOMAIN 964 1000

FT DOMAIN 1002 1040

FT DOMAIN 1042 1081

FT DOMAIN 1083 1122

FT DOMAIN 1126 1167

FT REPEAT 1168 1208

FT REPEAT 1209 1242

FT REPEAT 1243 1282

FT SIGNAL 1 20

FT CHAIN 21 1964

FT CHAIN 1411 1964

FT CHAIN 1428 1964

FT CHAIN 1463 1964

FT DOMAIN 21 1443

FT TRANSMEM 1444 1464

FT DOMAIN 1465 1964

FT DOMAIN 21 60

FT DOMAIN 61 112

FT DOMAIN 115 152

FT DOMAIN 153 189

FT DOMAIN 191 229

FT DOMAIN 231 271

FT DOMAIN 273 309

FT DOMAIN 311 350

FT DOMAIN 352 388

FT DOMAIN 389 427

FT DOMAIN 429 470

FT DOMAIN 472 508

FT DOMAIN 510 546

FT DOMAIN 548 584

FT DOMAIN 586 622

FT DOMAIN 623 656

FT DOMAIN 658 686

FT DOMAIN 688 724

FT DOMAIN 726 762

FT DOMAIN 764 800

FT DOMAIN 803 839

FT DOMAIN 841 877

FT DOMAIN 878 924

FT DOMAIN 926 962

FT DOMAIN 964 1000

FT DOMAIN 1002 1040

FT DOMAIN 1042 1081

FT DOMAIN 1083 1122

FT DOMAIN 1126 1167

FT REPEAT 1168 1208

FT REPEAT 1209 1242

FT REPEAT 1243 1282

FT SIGNAL 1 20

FT CHAIN 21 1964

FT CHAIN 1411 1964

FT CHAIN 1428 1964

FT CHAIN 1463 1964

FT DOMAIN 21 1443

FT TRANSMEM 1444 1464

FT DOMAIN 1465 1964

FT DOMAIN 21 60

FT DOMAIN 61 112

FT DOMAIN 115 152

FT DOMAIN 153 189

FT DOMAIN 191 229

FT DOMAIN 231 271

FT DOMAIN 273 309

FT DOMAIN 311 350

FT DOMAIN 352 388

FT DOMAIN 389 427

FT DOMAIN 429 470

FT DOMAIN 472 508

FT DOMAIN 510 546

FT DOMAIN 548 584

FT DOMAIN 586 622

FT DOMAIN 623 656

FT DOMAIN 658 686

FT DOMAIN 688 724

FT DOMAIN 726 762

FT DOMAIN 764 800

FT DOMAIN 803 839

FT DOMAIN 841 877

FT DOMAIN 878 924

FT DOMAIN 926 962

FT DOMAIN 964 1000

FT DOMAIN 1002 1040

FT DOMAIN 1042 1081

FT DOMAIN 1083 1122

FT DOMAIN 1126 1167

FT REPEAT 1168 1208

FT REPEAT 1209 1242

FT REPEAT 1243 1282

FT SIGNAL 1 20

FT CHAIN 21 1964

FT CHAIN 1411 1964

FT CHAIN 1428 1964

FT CHAIN 1463 1964

FT DOMAIN 21 1443

FT TRANSMEM 1444 1464

FT DOMAIN 1465 1964

FT DOMAIN 21 60

FT DOMAIN 61 112

FT DOMAIN 115 152

FT DOMAIN 153 189

FT DOMAIN 191 229

FT DOMAIN 231 271

FT DOMAIN 273 309

FT DOMAIN 311 350

FT DOMAIN 352 388

FT DOMAIN 389 427

FT DOMAIN 429 470

FT DOMAIN 472 508

FT DOMAIN 510 546

FT DOMAIN 548 584

FT DOMAIN 586 622

FT DOMAIN 623 656

FT DOMAIN 658 686

FT DOMAIN 688 724

FT DOMAIN 726 762

FT DOMAIN 764 800

FT DOMAIN 803 839

FT DOMAIN 841 877

FT DOMAIN 878 924

FT DOMAIN 926 962

FT DOMAIN 964 1000

FT DOMAIN 1002 1040

FT DOMAIN 1042 1081

FT DOMAIN 1083 1122

FT DOMAIN 1126 1167

FT REPEAT 1168 1208

FT REPEAT 1209 1242

FT REPEAT 1243 1282

FT SIGNAL 1 20

FT CHAIN 21 1964

FT CHAIN 1411 1964

FT CHAIN 1428 1964

FT CHAIN 1463 1964

FT DOMAIN 21 1443

FT TRANSMEM 1444 1464

FT DOMAIN 1465 1964

FT DOMAIN 21 60

FT DOMAIN 61 112

FT DOMAIN 115 152

FT DOMAIN 153 189

FT DOMAIN 191 229

FT DOMAIN 231 271

FT DOMAIN 273 309

FT DOMAIN 311 350

FT DOMAIN 352 388

FT DOMAIN 389 427

FT DOMAIN 429 470

FT DOMAIN 472 508

FT DOMAIN 510 546

FT DOMAIN 548 584

FT DOMAIN 586 622

FT DOMAIN 623 656

FT DOMAIN 658 686

FT DOMAIN 688 724

FT DOMAIN 726 762

FT DOMAIN 764 800

FT DOMAIN 803 839

FT DOMAIN 841 877

FT DOMAIN 878 924

FT DOMAIN 926 962

FT DOMAIN 964 1000

FT DOMAIN 1002 1040

FT DOMAIN 1042 1081

FT DOMAIN 1083 1122

FT DOMAIN 1126 1167

FT REPEAT 1168 1208

FT REPEAT 1209 1242

FT REPEAT 1243 1282

FT SIGNAL 1 20

FT CHAIN 21 1964

FT CHAIN 1411 1964

FT CHAIN 1428 1964

FT CHAIN 1463 1964

FT DOMAIN 21 1443

FT TRANSMEM 1444 1464

FT DOMAIN 1465 1964

FT DOMAIN 21 60

FT DOMAIN 61 112

FT DOMAIN 115 152

FT DOMAIN 153 189

FT DOMAIN 191 229

FT DOMAIN 231 271

FT DOMAIN 273 309

FT DOMAIN 311 350

FT DOMAIN 352 388

FT DOMAIN 389 427

FT DOMAIN 429 470

FT DOMAIN 472 508

FT DOMAIN 510 546

FT DOMAIN 548 584

FT DOMAIN 586 622

FT DOMAIN 623 656

FT DOMAIN 658 686

FT DOMAIN 688 724

FT DOMAIN 726 762

FT DOMAIN 764 800

FT DOMAIN 803 839

FT DOMAIN 841 877

FT DOMAIN 878 924

FT DOMAIN 926 962

FT DOMAIN 964 1000

FT DOMAIN 1002 1040

FT DOMAIN 1042 1081

FT DOMAIN 1083 1122

FT DOMAIN 1126 1167

FT REPEAT 1168 1208

FT REPEAT 1209 1242

FT REPEAT 1243 1282

FT SIGNAL 1 20

FT CHAIN 21 1964

FT CHAIN 1411 1964

FT CHAIN 1428 1964

FT CHAIN 1463 1964

FT DOMAIN 21 1443

FT TRANSMEM 1444 1464

FT DOMAIN 1465 1964

FT DOMAIN 21 60

FT DOMAIN 61 112

FT DOMAIN 115 152

FT DOMAIN 153 189

FT DOMAIN 191 229

FT DOMAIN 231 271

FT DOMAIN 273 309

FT DOMAIN 311 350

FT DOMAIN 352 388

FT DOMAIN 389 427

FT DOMAIN 429 470

FT DOMAIN 472 508

FT DOMAIN 510 546

FT DOMAIN 548 584

FT DOMAIN 586 622

FT DOMAIN 623 656

FT DOMAIN 658 686

FT DOMAIN 688 724

FT DOMAIN 726 762

FT DOMAIN 764 800

FT DOMAIN 803 839

FT DOMAIN 841 877

FT DOMAIN 878 924

FT DOMAIN 926 962

FT DOMAIN 964 1000

FT DOMAIN 1002 1040

FT DOMAIN 1042 1081

FT DOMAIN 1083 1122

FT DOMAIN 1126 1167

FT REPEAT 1168 1208

FT REPEAT 1209 1242

FT REPEAT 1243 1282

FT SIGNAL 1 20

FT CHAIN 21 1964

FT CHAIN 1411 1964

FT CHAIN 1428 1964

FT CHAIN 1463 1964

FT DOMAIN 21 1443

FT TRANSMEM 1444 1464

FT DOMAIN 1465 1964

FT DOMAIN 21 60

FT DOMAIN 61 112

FT DOMAIN 115 152

FT DOMAIN 153 189

FT DOMAIN 191 229

FT DOMAIN 231 271

FT DOMAIN 273 309

FT DOMAIN 311 350

FT DOMAIN 352 388

FT DOMAIN 389 427

FT DOMAIN 429 470

FT DOMAIN 472 508

FT DOMAIN 510 546

FT DOMAIN 548 584

FT DOMAIN 586 622

FT DOMAIN 623 656

FT DOMAIN 658 686

FT DOMAIN 688 724

FT DOMAIN 726 762

FT DOMAIN 764 800

FT DOMAIN 803 839

FT DOMAIN 841 877

FT DOMAIN 878 924

FT DOMAIN 926 962

FT DOMAIN 964 1000

FT DOMAIN 1002 1040

FT DOMAIN 1042 1081

FT DOMAIN 1083 1122

FT DOMAIN 1126 1167

FT REPEAT 1168 1208

FT REPEAT 1209 1242

FT REPEAT 1243 1282

FT SIGNAL 1 20

FT CHAIN 21 1964

FT CHAIN 1411 1964

FT CHAIN 1428 1964

FT CHAIN 1463 1964

FT DOMAIN 21 1443

FT TRANSMEM 1444 1464

FT DOMAIN 1465 1964

FT DOMAIN 21 60

FT DOMAIN 61 112

FT DOMAIN 115 152

FT DOMAIN 153 189

FT DOMAIN 191 229

FT DOMAIN 231 271

FT DOMAIN 273 309

FT DOMAIN 311 350

FT DOMAIN 352 388

FT DOMAIN 389 427

FT DOMAIN 429 470

FT DOMAIN 472 508

FT DOMAIN 510 546

FT DOMAIN 548 584

FT DOMAIN 586 622

FT DOMAIN 623 656

FT DOMAIN 658 686

FT DOMAIN 688 724

FT DOMAIN 726 762

FT DOMAIN 764 800

FT DOMAIN 803 839

FT DOMAIN 841 877

FT DOMAIN 878 924

FT DOMAIN 926 962

FT DOMAIN 964 1000

FT DOMAIN 1002 1040

FT DOMAIN 1042 1081

FT DOMAIN 1083 1122

FT DOMAIN 1126 1167

FT REPEAT 1168 1208

FT REPEAT 1209 1242

FT REPEAT 1243 1282

FT SIGNAL 1 20

FT CHAIN 21 1964

FT CHAIN 1411 1964

FT CHAIN 1428 1964

FT CHAIN 1463 1964

FT DOMAIN 21 1443

FT TRANSMEM 1444 1464

FT DOMAIN 1465 1964

FT DOMAIN 21 60

FT DOMAIN 61 112

FT DOMAIN 115 152

FT DOMAIN 153 189

FT DOMAIN 191 229

FT DOMAIN 231 271

FT DOMAIN 273 309

FT DOMAIN 311 350

FT DOMAIN 352 388

FT DOMAIN 389 427

FT DOMAIN 429 470

FT DOMAIN 472 508

FT DOMAIN 510 546

FT DOMAIN 548 584

FT DOMAIN 586 622

FT DOMAIN 623 656

FT DOMAIN 658 686

FT DOMAIN 688 724

FT DOMAIN 726 762

FT DOMAIN 764 800

FT DOMAIN 803 839

FT DOMAIN 841 877

FT DOMAIN 878 924

FT DOMAIN 926 962

FT DOMAIN 96

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 22, 2004, 12:42:05 ; Search time 18 seconds  
(without alignments)  
789.731 Million cell updates/sec

Title: US-09-978-191A-506

Perfect score: 1505

Sequence: 1 MRGSEVLLMVLVAVGGT.....SEQISFLEQLGSCCKKDS 273

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_42.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	196	13.0	1964	1 NTC4_MOUSE	P31695 mus musculus
2	188.5	12.5	2003	1 NTC4_MOUSE	Q99466 homo sapien
3	185	12.3	816	1 NEL2_MOUSE	Q61220 mus musculus
4	183.5	12.2	2907	1 FBN2_MOUSE	Q61555 mus musculus
5	182.5	12.1	816	1 NEL2_MOUSE	Q99435 homo sapien
6	182	12.0	810	1 NEL1_HUMAN	Q92832 homo sapien
7	180.5	12.0	816	1 NEL2_RAT	Q62918 rattus norv
8	180	12.0	810	1 NEL1_RAT	Q62919 rattus norv
9	176.5	11.7	816	1 NEL1_CHICK	Q30827 gallus gall
10	176.5	11.7	2871	1 FBN1_HUMAN	P35555 homo sapien
11	175.5	11.7	2321	1 NTC3_HUMAN	Q9um47 homo sapien
12	175	11.6	652	1 CD93_HUMAN	Q9np73 homo sapien
13	175	11.6	2703	1 NOTC_DROME	P07207 drosophila
14	173.5	11.5	1247	1 NID0_HUMAN	P14543 homo sapien
15	172.5	11.5	2871	1 FBN1_MOUSE	Q61554 mus musculus
16	172	11.4	2531	1 NTC1_MOUSE	Q01705 mus musculus
17	171.5	11.4	2871	1 FBN1_BOVIN	P38133 bos taurus
18	170.5	11.3	2871	1 FBN1_PIG	Q9tvc6 sus scrofa
19	170	11.3	675	1 FRTS_MOUSE	Q08761 mus musculus
20	170	11.3	2911	1 FBN2_HUMAN	P35556 homo sapien
21	169.5	11.3	2318	1 NTC3_MOUSE	Q61982 mus musculus
22	169	11.2	833	1 DL_DROME	P10041 drosophila
23	168	11.2	675	1 PRTS_RAT	P53813 rattus norv
24	165	11.0	1238	1 JAG2_HUMAN	Q9y219 homo sapien
25	165	11.0	1247	1 JAG2_MOUSE	Q9y65 mus musculus
26	164	10.9	1202	1 JAG2_RAT	P97607 rattus norv
27	164	10.9	1429	1 L112_CABEL	P14585 caenorhabdi
28	164	10.9	2524	1 NOTC_XENLA	P21783 xenopus lae
29	163.5	10.9	2556	1 NTC1_HUMAN	P46531 homo sapien
30	163	10.8	644	1 CD93_MOUSE	Q89103 mus musculus
31	162.5	10.8	2319	1 NTC3_RAT	Q9x172 rattus norv
32	162	10.8	459	1 PRTC_PIG	Q9gip2 sus scrofa
33	162	10.8	646	1 FRTS_RABIT	P98118 oryctolagus

## ALIGNMENTS

### RESULT 1

#### NTC4\_MOUSE

ID NTC4\_MOUSE STANDARD; PRT; 1964 AA.  
AC P31695; Q35442; O88314; O88316; Q62389; Q62390; Q9RLW9; Q9RLX0;  
DT 01-JUN-1993 (Rel. 26, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Neogenic locus notch homolog protein 4 precursor (Notch 4)  
DE [Contains: Transforming protein Int-3].  
GN NOTCH4 OR INT3 OR INT-3.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=92194507; PubMed=1312643;  
RA Robbins J., Blondel B.J., Gallahan D., Callahan R.;  
RT "Mouse mammary tumor gene Int-3: a member of the notch gene family  
RT transforms mammary epithelial cells.";  
RL J. Virol. 66:2594-2599(1992).  
RN [2]  
RP REVISIONS, SEQUENCE FROM N.A.  
RX MEDLINE=97294599; PubMed=9150355;  
RA Gallahan D., Callahan R.;  
RT "The mouse mammary tumor associated gene INT3 is a unique member of  
RT the NOTCH gene family (NOTCH4).";  
RL Oncogene 14:1883-1890(1997).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Lung, and Testis;  
RX MEDLINE=96281668; PubMed=8681805;  
RA Uytendaele H., Marazzi G., Wu G., Yan Q., Sassoon D., Kitajewski J.;  
RT "Notch4/int-3, a mammary proto-oncogene, is an endothelial  
RT cell-specific mammalian Notch gene.";  
RL Development 122:2251-2259(1996).  
RN [4]  
RP SEQUENCE FROM N.A.  
RA Rowen L., Mahairas G., Qin S., Ahearn M.E., Dankers C., Lasky S.,  
RA Loretz C., Schmidt S., Tipton S., Traicoff R., Zackrone K., Hood L.;  
RT "Sequence of the mouse major histocompatibility locus class III  
RT region.";  
RL Submitted (Oct-1997) to the EMBL/GenBank/DBJ databases.  
RN [5]  
RP SEQUENCE OF 1436-1600 FROM N.A.  
RX MEDLINE=9925212; PubMed=10233982;  
RA Lee J.-S., Haruna T., Ishimoto A., Honjo T., Yanagawa S.-I.;  
RT "Intracisternal type A particle-mediated activation of the Notch4/int3  
RT gene in a mouse mammary tumor: generation of truncated Notch4/int3  
RT mRNAs by retroviral splicing events.";  
RL J. Virol. 73:5166-5171(1999).  
RN [6]  
RP FUNCTION  
RX MEDLINE=21244657; PubMed=11344305;  
RA Uytendaele H., Ho J., Rossant J., Kitajewski J.;  
RT "Vascular patterning defects associated with expression of activated

34 162 10.8 833 1 SRC2\_MOUSE  
35 162 10.8 870 1 SRC2\_HUMAN  
36 161 10.7 675 1 PRTS\_BOVIN  
37 160.5 10.7 618 1 DLL3\_HUMAN  
38 160.5 10.7 2531 1 NTC1\_RAT  
39 160 10.6 2437 1 NTC1\_BRARE  
40 159.5 10.6 643 1 CD93\_RAT  
41 158.5 10.5 407 1 FA7\_BOVIN  
42 158.5 10.5 456 1 PRTC\_BOVIN  
43 157.5 10.5 379 1 WIF1\_MOUSE  
44 157 10.4 676 1 PRTS\_HUMAN  
45 156.5 10.4 1213 1 JAG3\_BRARE  
P59222 mus musculus  
Q96966 homo sapien  
P07324 bos taurus  
Q9ny17 homo sapien  
Q07008 rattus norv  
P46530 brachydanio  
Q9et61 rattus norv  
P22457 bos taurus  
P00745 bos taurus  
Q9wua1 mus musculus  
P07225 homo sapien  
Q90y54 brachydanio



```
Db 43 VCAEQKTLVGHROPCTVQAFSRIVPVWVRRTGCAQAQAWCIGQERTTVYMSYRQVYA-TEA 101
Qy 83 RPRYACCPGKMTSGLECA-----CGAAI--COPPCRN--GGSCVQPGRCRCPAGWR-- 130
Db 102 RTVFRCCPGWSQKPGQGLSDVDECAASAGGCGPCNTVGGF-----YCRCPGGVQLQ 156
Qy 131 --GPTCQSDVDECSARGGCPQRCINTAGSYWCOCWEGHLSADGTLCV 177
Db 157 GDGKTCQ-DVDECAHNGGQHCRCVNTPGSYLCCKPGLHTDGRCL 204

RESULT 12
Q9UFK6 PRELIMINARY; PRT; 558 AA.
AC Q9UFK6
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein (Fragment).
GN DKZP564P2063.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Dueterhoeft A.; Lauber J.; Mewes H.W.; Gassenhuber J.; Wiemann S.;
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 1 MAM DOMAIN.
DR EMBL; AL117610; CAB56014.1; -.
DR PIR; T17324; T17324.
DR HSP; P00736; IAPQ.
DR Genew; HGNC:3235; EGFL6.
DR GO; GO:0005615; C:extracellular space; TAS.
DR GO; GO:0005178; F:integrin binding; TAS.
DR GO; GO:0007049; P:cell cycle; TAS.
DR InterPro; IPR000152; ASX_HYDROXYL_S.
DR InterPro; IPR001881; EGF_CA.
DR InterPro; IPR006209; EGF_Like.
DR InterPro; IPR000998; MAM_domain.
DR Pfam; PF00008; EGF; 4.
DR Pfam; PF00629; MAM; 1.
DR SMART; SM00179; EGF_CA; 3.
DR PROSITE; PS00010; ASX_HYDROXYL; 3.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; 2.
DR PROSITE; PS01187; EGF_CA; 2.
DR PROSITE; PS50060; MAM_2; 1.
KW Hypothetical protein; EGF-like domain; Glycoprotein.
FT NON_TER
SQ SEQUENCE 558 AA; 61828 MW; AA38D7DCE402BFA3 CRC64;

Query Match 16.8%; Score 252.5; DB 4; Length 558;
Best Local Similarity 36.8%; Pred. No. 2.6e-14;
Matches 57; Conservative 17; Mismatches 50; Indels 31; Gaps 7;

Qy 58 GHRACSTYRTYRTAYRSPGLAPRPRYACCPGKMTSGLECAQAACIQQPCRNNGSC 117
Db 26 GNAASARHHGLASA--RQPGVCHYGTGLACCYGRNRNS--KGVC-EATCEPGCK-FGEC 79
Qy 118 VOPGRCRCPAGWRGDTCSQDVDECSARGGCPQRCINTAGSYWCOCWEGHLSADGTLCV 177
Db 80 VGNKRCCLPGYTGKTCSDVNECGMKPRPCQHCRCVNTGHSYKFCFLSGHMLPDT-CV 138
Qy 178 -----PKGPP-----RVAPN 188
Db 139 YSRTCAMINCQYSCDTEEGPQCLCPSSGLRLAPN 173

RESULT 13
Q9NZL7 PRELIMINARY; PRT; 553 AA.
ID Q9NZL7
Qy 58 GHRACSTYRTYRTAYRSPGLAPRPRYACCPGKMTSGLECAQAACIQQPCRNNGSC 117
Db 26 GNAASARHHGLASA--RQPGVCHYGTGLACCYGRNRNS--KGVC-EATCEPGCK-FGEC 79
Qy 118 VOPGRCRCPAGWRGDTCSQDVDECSARGGCPQRCINTAGSYWCOCWEGHLSADGTLCV 177
Db 80 VGNKRCCLPGYTGKTCSDVNECGMKPRPCQHCRCVNTGHSYKFCFLSGHMLPDT-CV 138
Qy 178 -----PKGPP-----RVAPN 188
Db 139 YSRTCAMINCQYSCDTEEGPQCLCPSSGLRLAPN 173
```

```
Q9NZL7;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Epidermal growth factor repeat containing protein.
GN EGF16.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20079166; PubMed=10610727;
RA Yeung G.; Muller J.J.; Berntsen R.P.; Loeb D.B.; Drmanac R.;
RA Ford J.E.;
RT "Cloning of a novel epidermal growth factor repeat containing gene
RT EGF16; expressed in tumor and fetal tissues.";
RL Genomics 62:304-307(1999).
CC -1- SIMILARITY: CONTAINS 1 MAM DOMAIN.
DR EMBL; AF186084; AAF27812.1; -.
DR HSP; P00736; IAPQ.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR InterPro; IPR000152; ASX_HYDROXYL_S.
DR InterPro; IPR001881; EGF_CA.
DR InterPro; IPR006209; EGF_Like.
DR InterPro; IPR000998; MAM_domain.
DR Pfam; PF00008; EGF; 4.
DR Pfam; PF00629; MAM; 1.
DR SMART; SM00179; EGF_CA; 3.
DR PROSITE; PS00010; ASX_HYDROXYL; 3.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; 2.
DR PROSITE; PS01187; EGF_CA; 3.
DR PROSITE; PS50060; MAM_2; 1.
KW EGF-like domain; Glycoprotein.
SQ SEQUENCE 553 AA; 61314 MW; 2PF55F167857DE50 CRC64;

Query Match 16.7%; Score 251.5; DB 4; Length 553;
Best Local Similarity 36.8%; Pred. No. 3.1e-14;
Matches 57; Conservative 17; Mismatches 50; Indels 31; Gaps 7;

Qy 58 GHRACSTYRTYRTAYRSPGLAPRPRYACCPGKMTSGLECAQAACIQQPCRNNGSC 117
Db 21 GNAASARHHGLASA--RQPGVCHYGTGLACCYGRNRNS--KGVC-EATCEPGCK-FGEC 74
Qy 118 VOPGRCRCPAGWRGDTCSQDVDECSARGGCPQRCINTAGSYWCOCWEGHLSADGTLCV 177
Db 75 VGNKRCCLPGYTGKTCSDVNECGMKPRPCQHCRCVNTGHSYKFCFLSGHMLPDT-CV 133
Qy 178 -----PKGPP-----RVAPN 188
Db 134 NSRTCAMINCQYSCDTEEGPQCLCPSSGLRLAPN 168

RESULT 14
Q8IUX8 PRELIMINARY; PRT; 553 AA.
ID Q8IUX8
AC Q8IUX8
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Similar to EGF-like domain, multiple 6.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Strausberg R.;
RL Submitted (OCT-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC038587; AAH38587.1; -.

```





RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,  
RA Diatchenko L., Marusina K., Rubin A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blackley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
RA Krzywinski M.I., Skalska U., Smalhus D.E., Schnerch A., Schein J.E.,  
RA Jones S.J., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences";  
EL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Skin;  
RA Strausberg R.;  
RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.  
DR EMBL; U89336; AAB47494.1; -;  
DR EMBL; BC052591; AAB52591.1; -;  
DR HSRF; P35555; IENM.  
DR GO; GO:0005509; F:calcium ion binding; IEA.  
DR InterPro; IPR00152; Asx hydroxylase.  
DR InterPro; IPR001881; EGF Ca.  
DR InterPro; IPR006209; EGF-like.  
DR Pfam; PF00008; EGF; 2.  
DR SMART; SM00179; EGF CA; 1.  
DR PROSITE; PS00010; ASX HYDROXYL; 1.  
DR PROSITE; PS00022; EGF 1; 1.  
DR PROSITE; PS01186; EGF 2; 1.  
DR PROSITE; PS01187; EGF CA; 1.  
KW EGF-like domain.  
SQ SEQUENCE 293 AA; 32262 MW; 6519CA254FB23FD0 CRC64;  
  
Query Match 28.1%; Score 423.5; DB 4; Length 293;  
Best Local Similarity 36.5%; Pred. No. 5e-30;  
Matches 103; Conservative 37; Mismatches 107; Indels 35; Gaps 10;  
  
QY 12 LLVLAVGTEHAYRPGRRVCAVRAHGDV--SESFVQRYVQPLTTCGHRACSTYRTLY 69  
DB 19 LLIPGEGAGKGSRLRESQVCSKQTLVPLVPLHYNESYSQPVYKPYLTLCAGRRICSTYRTMY 78  
  
QY 70 RTAYRSPGLAPAPRYACCPGWKRTSGLPGA--CGAAICOPPCRNNGSCVQPGRCRCPA 127  
DB 79 RVWM-REVREVVQTHAVCCQGWKKRH--PGALTC-EAICAKPCLNGVCVRPDQCECAP 134  
  
QY 128 GWRGDTCSQDVDECSARRGGCPQRCINTAGSYWCOCWEHSLSDGTLCPVKGPPRVAP 187  
DB 135 GWGKGKCHVDVDECRITSLCHHCFNTAGSFTCGCPHDVLVLDGRTCMGSPPEPTSA 194  
  
QY 188 NPTGV-----DSAMKEEVORLSRVLDLLEKQLVLAHLSASA---LEHGLP-D 235  
DB 195 SILSVAVREAEKDERALKQEIHELGRLEQLQ-----WAGQAGAVRAVLVPP 243  
  
QY 236 PGSLLVHVSFOQL-----GRIDSLSQISFLERQLGSCSKKDS 273  
DB 244 PEELQPEQVAVELWGRDRIEQLSDQVLLLEERLGACSCEDNS 285  
  
RESULT 9  
Q81V30  
ID Q81V30 PRELIMINARY; PRT; 293 AA.  
AC Q81V30;  
DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
DR 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Chromosome 6 open reading frame 8.  
OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Brain;  
RA Strausberg R.;  
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC035574; AAB35574.1; -;  
DR GO; GO:0005509; F:calcium ion binding; IEA.  
DR InterPro; IPR000152; Asx hydroxylase.  
DR InterPro; IPR001881; EGF Ca.  
DR InterPro; IPR006209; EGF-like.  
DR InterPro; IPR006210; IEGF.  
DR Pfam; PF00008; EGF; 2.  
DR SMART; SM00181; EGF; 2.  
DR SMART; SM00179; EGF CA; 1.  
DR PROSITE; PS00010; ASX HYDROXYL; 1.  
DR PROSITE; PS00022; EGF 1; 1.  
DR PROSITE; PS01186; EGF 2; 1.  
DR PROSITE; PS01187; EGF CA; 1.  
SQ SEQUENCE 293 AA; 32261 MW; 6519CA255568FFD0 CRC64;  
  
Query Match 27.9%; Score 420.5; DB 4; Length 293;  
Best Local Similarity 36.2%; Pred. No. 9.4e-30;  
Matches 102; Conservative 38; Mismatches 107; Indels 35; Gaps 10;  
  
QY 12 LLVLAVGTEHAYRPGRRVCAVRAHGDV--SESFVQRYVQPLTTCGHRACSTYRTLY 69  
DB 19 LLIPGEGAGKGSRLRESQVCSKQTLVPLVPLHYNESYSQPVYKPYLTLCAGRRICSTYRTMY 78  
  
QY 70 RTAYRSPGLAPAPRYACCPGWKRTSGLPGA--CGAAICOPPCRNNGSCVQPGRCRCPA 127  
DB 79 RVWM-REVREVVQTHAVCCQGWKKRH--PGALTC-EAICAKPCLNGVCVRPDQCECAP 134  
  
QY 128 GWRGDTCSQDVDECSARRGGCPQRCINTAGSYWCOCWEHSLSDGTLCPVKGPPRVAP 187  
DB 135 GWGKGKCHVDVDECRITSLCHHCFNTAGSFTCGCPHDVLVLDGRTCMGSPPEPTSA 194  
  
QY 188 NPTGV-----DSAMKEEVORLSRVLDLLEKQLVLAHLSASA---LEHGLP-D 235  
DB 195 SILSVAVREAEKDERALKQEIHELGRLEQLQ-----WAGQAGAVRAVLVPP 243  
  
QY 236 PGSLLVHVSFOQL-----GRIDSLSQISFLERQLGSCSKKDS 273  
DB 244 PEELQPEQVAVELWGRDRIEQLSDQVLLLEERLGACSCEDNS 285  
  
RESULT 10  
Q95RQ1  
ID Q95RQ1 PRELIMINARY; PRT; 512 AA.  
AC Q95RQ1; Q9VZD0;  
DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE LD16414P (CG7447-PA) (CG7447-PA).  
GN CG7447.  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=7227;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Berkley;  
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,  
RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,  
RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,  
RA Nuncio J., Pacleb J., Paragas V., Park S., Phouanavong S., Wan K.,  
RA Yu C., Lewis S.E., Rubin G.M., Celisner S.;  
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.



```
Db 141 LASRSPHGLQDPGSLLAHSFQQLDRIDSLSEQVFLSEQLGSCCKD 189

RESULT 6
Q7ZXTO PRELIMINARY; PRT; 280 AA.
AC Q7ZXTO;
DT 01-JUN-2003 (TREMBlrel. 24, Created)
DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Similar to NEU1 protein.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RA Klein S., Strausberg R.;
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC044267; AAH4267.1; -.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR InterPro; IPR000152; Asx hydroxyl_S.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR006209; EGF_like.
DR Pfam; PF00008; EGF; 2.
DR SMART; SM00179; EGF_CA; 2.
DR PROSITE; PS00010; ASX HYDROXYL; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS01187; EGF_CA; 1.
DR Hypothetical protein; EGF-like domain.
KW SEQUENCE 280 AA; 30720 MW; 424EP8CF1B16D220 CRC64;

Query Match 45.3%; Score 682.5; DB 13; Length 280;
Best Local Similarity 48.7%; Pred. No. 2.2e-53;
Matches 132; Conservative 44; Mismatches 84; Indels 11; Gaps 5;

QY 11 MLLVLAV--GGTEHYRGRVCAVRAHGDPV--ESFVQRYVQFLLTCGHRACSTYR 66
DB 11 YLLILAVTSAADHLRYTGRICSDAGHPGVTSVTSQFVQVHSPIMTLCGHRICSTYR 70
QY 67 TYRTAYRSPCLAPARYACCPGKRTSLPGACGAAICQPPCRNGSCVQPCRCPC 126
DB 71 TTYKVSYQ--VSRKTSFPLSCCPGWRRAIGQTHSCGQALCRLOCGGTVCVSNKCEP 129
QY 127 AGWRGDTQSDVDECSARRGCGPQRCINTAGSYWCQWEGHSLSDGTLG--VPKGGPPR 184
DB 130 AGWRGTHCQMDVDECSGTHQCSQACINSAGSFCECLEGYRLMADGKTCRKVPAPTVP 189
QY 185 VAP---NPTGSDSAMKEEVQRLQSRVLDLEELKQLVLAPLHSLASQALEHGLPDPGSLV 241
DB 190 ASPTVSQSGIPSHSVKNEALRSKIDVLEQKHLHLLTFPQGLTFSPDDA-ADPIALLT 248
QY 242 HSFOQLGRIDSLSEQIFLEQLGSCCKD 272
DB 249 RSLQQLDRIDSLSEQIFLEERLETCSCKE 279

RESULT 7
Q35447 PRELIMINARY; PRT; 293 AA.
AC Q35447;
DT 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Hypothetical protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;

RN [1]
RP SEQUENCE FROM N.A.
RA Rowen L., Mahairas G., Qin S., Ahearn M.E., Dankers C., Lasky S.,
RA Loretz C., Schmidt S., Tipton S., Traicoff R., Zackrone K., Hood L.;
RT "Sequence of the mouse major histocompatibility locus class III
RT region.";
RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF030001; AAB82010.1; -.
DR PIR; T09065; T09065.
DR HSSP; P35555; LEWN.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR InterPro; IPR000152; Asx hydroxyl_S.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR006209; EGF_like.
DR Pfam; PF00008; EGF; 2.
DR SMART; SM00179; EGF_CA; 1.
DR PROSITE; PS00010; ASX HYDROXYL; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS01187; EGF_CA; 1.
DR Hypothetical protein; EGF-like domain.
KW SEQUENCE 293 AA; 32066 MW; 8A99ASEC000A2C62 CRC64;

Query Match 28.9%; Score 435.5; DB 11; Length 293;
Best Local Similarity 37.3%; Pred. No. 4.2e-31;
Matches 112; Conservative 36; Mismatches 101; Indels 51; Gaps 14;

QY 1 MERGSEVLLMVLVLAVG--GTE-HAYRGRVCAVRAHGDPV--SESFVQRYVQFLLTTC 56
DB 11 LRG-----LSFFVLVMTGSTRGSGKESLGVCSTQLLPLVLYNESQPYKYFLTLC 65
QY 57 DGHACSTYRTTYRTAYRSPCLAPARYACCPGKRTSLPGACGA-AICQPPCRNGG 115
DB 66 AGRICSTYRTTYRYVAVREVREVP-QTHVVCQCKKPH--PGALTCDALCKECLNGG 122
QY 116 SCVQPCRCPCAGWRGDTQSDVDECSARRGCGPQRCINTAGSYWCQWEGHSLSDGTL 175
DB 123 VCTGPDRCECAPGWSGKHCHVDVDECRASLTLCSHGCLNTLGSFLCSFHPVLVLDGRT 182
QY 176 CVPKGGPPRVANPTGV-----DSAMKEEVQRLQSRVLDLEELKQLVLAPL 221
DB 183 CA--GGPPE---SPTASLSLVAVREADSEERALARWEVAELGRLEKLEQ----- 228
QY 222 HSLASQA---LEHGLP-DFGSLLVHSFQOL-----GRDLSLSEIQIFLEQLGSCCKDS 273
DB 229 --WATQAGAVRAVLPMPEELRPEQVAELWGRGDRIESLSQVLLLEERLGACACEDNS 286

RESULT 8
Q99944 PRELIMINARY; PRT; 293 AA.
AC Q99944;
DT 01-MAY-1997 (TREMBlrel. 03, Created)
DT 01-MAY-1997 (TREMBlrel. 03, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE NG3 (C6orf8 protein).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Rowen L., Dankers C., Baskin D., Faust J., Loretz C., Ahearn M.E.,
RA Banta A., Spies T., Hood L.;
RT "Sequence determination of 300 kilobases of the human class III MHC
RT locus.";
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Skin;
RX MEDLINE=223388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
```

```

Db 64 RACSTYRTIYRTAYRRSPGVTPARPRYACCPGKRTSGLPGACAAICQPPCGNGGSCIR 123
QY 120 PGRCPAGWGTQSDVDECSARRGCGPCRCINTAGSYWCQWEGHSLSDGTLCPVK 179
Db 124 PGRCPAGWGTQSDVDECSARRGCGPCRCINTAGSYWCQWEGHSLSDGTLCPVK 183
QY 180 GGPSPVAPNPT-GVDSAMKEEVORLQSRVDLLEKQLVLAPLHSLASQALEHGLPDPGS 238
Db 184 EGSPVAPNPTAGVDSMAREEVYRLQARVDVLEKQLVLAPLHSLASRSTHGLQDPGS 243
QY 239 LHVHFOQLGRIDSLSEQISFLEHGLSCCKD 272
Db 244 LLAHSFOQLGRIDSLSEQISFLEHGLSCCKD 277

```

## RESULT 4

```

Q9DCP5 ID Q9DCP5 PRELIMINARY; PRT; 265 AA.
AC Q9DCP5;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Vascular endothelial zinc finger 1.
GN VFZF1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC SFRAIN=C57BL/6J; TISSUE=Kidney;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaudo I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staebli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
EL Nature 403:685-690(2001)
DR EMBL; AK002601; BAB22222.1; -.
DR HSSP; P00740; IEDM.
DR MGD; MGI:1313291; Vezfl.
DR GO; GO:0005509; P:calcium ion binding; IEA.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR006210; ISGF.
DR Pfam; PF00008; EGF; 1.
DR SMART; SM00181; EGF; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS01187; EGF_CA; 1.
DR EGF-like domain.
KW SEQUENCE 265 AA; 28665 MW; D81EC3DA884FF97E CRC64;

```

Query Match 73.3%; Score 1103.5; DB 11; Length 265;  
 Best Local Similarity 73.7%; Pred. No. 2.6e-91;  
 Matches 202; Conservative 18; Mismatches 39; Indels 15; Gaps 3;

QY 1 MRGSEVLLMLLVLAVGG-TEHAVRPGRRVCAVRAHGDVPVSEFVQRTVPTLTCDGH 59

```

Db 4 MWGSELIVANFLVLAADGTEHYRPSRRYCTVIGISGSSISFTFVQRTVPTLTCDGH 63
QY 60 RACSTYRTIYRTAYRRSPGLAPARPRYACCPGKRTSGLPGACAAICQPPCGNGGSCVQ 119
Db 64 RACSTYRTIYRTAYRRSPGVTPARPRYACCPGKRTSGLPGACAAICQPPCGNGGSCIR 123
QY 120 PGRCPAGWGTQSDVDECSARRGCGPCRCINTAGSYWCQWEGHSLSDGTLCPVK 179
Db 124 PGRCPAGWGTQSDVDECSARRGCGPCRCINTAGSYWCQWEGHSLSDGTLCPVK 183
QY 180 GGPSPVAPNPT-GVDSAMKEEVORLQSRVDLLEKQLVLAPLHSLASQALEHGLPDPGS 238
Db 184 EGSPVAPNPTAGVDSMAREEVYRLQARVDVLEKQLVLAPLHSLASRSTHGLQDPGS 243
QY 239 LHVHFOQLGRIDSLSEQISFLEHGLSCCKD 272
Db 244 LLAHSFOQLGRIDSLSEQISFLEHGLSCCKD 264

```

## RESULT 5

```

Q9JKW3 ID Q9JKW3 PRELIMINARY; PRT; 190 AA.
AC Q9JKW3;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE CEL20.
GN CEL20.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Spleen;
RX Chan M.T.W., Ng C.C.Y., Lim E.K.B., Huynh H.T.;
RA "Cloning and Characterization of a Novel 20.4kD Estrogen-regulated
RT protein in the Rat Spleen.";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Spleen;
RX Marcantonio D., Huynh H.T.;
RA Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
RL EMBL; AP223678; AAF35352.1; -.
DR GO; GO:0005509; P:calcium ion binding; IEA.
DR InterPro; IPR000152; ASx hydroxyl_S.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR006209; EGF_like.
DR Pfam; PF00008; EGF; 2.
DR SMART; SM00179; EGF_CA; 1.
DR PROSITE; PS00010; ASX HYDROXYL; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS01187; EGF_CA; 1.
DR EGF-like domain.
KW SEQUENCE 190 AA; 20527 MW; C540EF0687F1E998 CRC64;

```

Query Match 48.3%; Score 727.5; DB 11; Length 190;  
 Best Local Similarity 76.9%; Pred. No. 1.3e-57;  
 Matches 130; Conservative 15; Mismatches 23; Indels 1; Gaps 1;

```

QY 105 AICQPPCGNGGSCVQFGRRCRCPCAGWRGDTCSQDVDECSARRGCGPCRCINTAGSYWCQW 164
Db 21 AICQPPCGNGGSCVQFGRRCRCPCAGWRGDTCSQDVDECSARRGCGPCRCINTAGSYWCQW 80
QY 165 EGHSLSDGTLCPVKGGPPRVPAPNPT-GVDSAMKEEVORLQSRVDLLEKQLVLAPLHSL 223
Db 81 EQGSFSDADVLCLPKEGSPVAPSPFTFGVDSVRRVYVYKLQARVDVLEKQLVLAPLHSL 140
QY 224 LASQALEHGLPDPGSLVHSPQQLGRIDSLSEQISFLEHGLSCCKD 272

```

QY 61 ACSTYRTYRTAYRRSPGLAPRRYACCPGKRTSGLPGAGAAICQPPCRNGSCVQP 120  
 Db 61 ACSTYRTYRTAYRRSPGLAPRRYACCPGKRTSGLPGAGAAICQPPCRNGSCVQP 120  
 QY 121 GRCPCPAGWGDTCCSDVDECSARRGGCPQRCINTAGSYWCQWEGHSLSDGTLCPVKG 180  
 Db 121 GRCPCPAGWGDTCCSDVDECSARRGGCPQRCINTAGSYWCQWEGHSLSDGTLCPVKG 180

QY 181 GPRVAPNPTGVDSSAMKEEVQRLQSRVDLLEKQLVLAPLHSLASQALEHGLPDPGSL 240  
 Db 181 GPRVAPNPTGVDSSAMKEEVQRLQSRVDLLEKQLVLAPLHSLASQALEHGLPDPGSL 240

QY 241 VHSFOOLGRIDSLSQISFLEQLGSCCKKDS 273  
 Db 241 VHSFOOLGRIDSLSQISFLEQLGSCCKKDS 273

## RESULT 2

Q9UHF1 PRELIMINARY; PRT; 273 AA.  
 AC Q9UHF1;  
 DT 01-MAY-2000 (TREMELrel. 13, Created)  
 DT 01-MAY-2000 (TREMELrel. 13, Last sequence update)  
 DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)  
 DE NOTCH4-like protein (Hypothetical protein).  
 GN ZNEU1 OR DRKZF586L2317.  
 OS Homo sapiens (human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Sheppard P., Jelinek L., Whitmore T., Blumberg H., Lehner J.,  
 RA O'Hara P.;  
 RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC Tissue=Uterus;  
 RA Koehrer K., Beyer A., Mewes H.W., Gassenhuber J., Wiemann S.;  
 RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF186111; AAF01429.1; -;  
 DR EMBL; AL512735; CAC21666.1; -;  
 DR HSSP; P08709; 1PAX.  
 DR GO; GO:0005509; F:calcium ion binding; IEA.  
 DR InterPro; IPR000152; Asx hydroxyl\_S.  
 DR InterPro; IPR001881; EGF\_Ca.  
 DR InterPro; IPR006209; EGF\_Like.  
 DR Pfam; PF00008; EGF\_2.  
 DR SMART; SM00179; EGF\_CA; 1.  
 DR PROSITE; PS00010; ASX HYDROXYL; 1.  
 DR PROSITE; PS00022; EGF\_1; 1.  
 DR PROSITE; PS01186; EGF\_2; 1.  
 DR PROSITE; PS01187; EGF\_CA; 1.  
 DR Hypothetical protein; "EGF-like domain."  
 KW SEQUENCE 273 AA; 23617 MW; 5740B845ED5A988 CRC64;

Query Match 99.9%; Score 1504; DB 4; Length 273;  
 Best Local Similarity 99.6%; Pred. No. 2.2e-127;  
 Matches 272; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRGSEVLLMWLLVAVGGTEHAYRPGRRYACCPGKRTSGLPGAGAAICQPPCRNGSCVQP 60  
 Db 1 MRGSEVLLMWLLVAVGGTEHAYRPGRRYACCPGKRTSGLPGAGAAICQPPCRNGSCVQP 60  
 QY 61 ACSTYRTYRTAYRRSPGLAPRRYACCPGKRTSGLPGAGAAICQPPCRNGSCVQP 120  
 Db 61 ACSTYRTYRTAYRRSPGLAPRRYACCPGKRTSGLPGAGAAICQPPCRNGSCVQP 120  
 QY 121 GRCPCPAGWGDTCCSDVDECSARRGGCPQRCINTAGSYWCQWEGHSLSDGTLCPVKG 180  
 Db 121 GRCPCPAGWGDTCCSDVDECSARRGGCPQRCINTAGSYWCQWEGHSLSDGTLCPVKG 180

QY 181 GPRVAPNPTGVDSSAMKEEVQRLQSRVDLLEKQLVLAPLHSLASQALEHGLPDPGSL 240  
 Db 181 GPRVAPNPTGVDSSAMKEEVQRLQSRVDLLEKQLVLAPLHSLASQALEHGLPDPGSL 240  
 QY 241 VHSFOOLGRIDSLSQISFLEQLGSCCKKDS 273  
 Db 241 VHSFOOLGRIDSLSQISFLEQLGSCCKKDS 273

## RESULT 3

Q9OXT5 PRELIMINARY; PRT; 278 AA.  
 AC Q9OXT5;  
 DT 01-MAY-2000 (TREMELrel. 13, Created)  
 DT 01-MAY-2000 (TREMELrel. 13, Last sequence update)  
 DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)  
 DE NOTCH4-like protein (Vascular endothelial zinc finger 1) (EGF-like domain 7).  
 GN EGF17 OR VEZF1 OR ZNEU1.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Sheppard P., Jelinek L., Whitmore T., Blumberg H., Lehner J.,  
 RA O'Hara P.;  
 RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Presnell S., Gilbert T., Whitmore T., Foster D., Hart C., Lehner J.,  
 RA Martinez T., Hoffmann R., O'Hara P.;  
 RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC Tissue=Liver;  
 RA Strausberg R.;  
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CD-1;  
 RA Fitch M.J., Kuhnert F., Stuhlmann H.;  
 RL Isolation of an early murine embryonic gene, Egf17.;  
 DR EMBL; AF184973; AAF01322.1; -;  
 DR EMBL; BC024610; AAH24610.1; -;  
 DR EMBL; AY309459; AAP74732.1; -;  
 DR HSSP; P00740; 1EDM.  
 DR MGD; MGI:1313291; Vezfl.  
 DR MGD; MGI:2449923; Egf17.  
 DR GO; GO:0005509; F:calcium ion binding; IEA.  
 DR InterPro; IPR001881; EGF\_Ca.  
 DR InterPro; IPR006209; EGF\_Like.  
 DR InterPro; IPR006210; IEGF.  
 DR Pfam; PF00008; EGF\_1.  
 DR SMART; SM00181; EGF\_1.  
 DR PROSITE; PS00022; EGF\_1; 1.  
 DR PROSITE; PS01186; EGF\_2; 1.  
 DR PROSITE; PS01187; EGF\_CA; 1.  
 KW EGF-like domain.  
 SQ SEQUENCE 278 AA; 30125 MW; 863159F567102FA7 CRC64;

Query Match 78.4%; Score 1180; DB 11; Length 278;  
 Best Local Similarity 77.7%; Pred. No. 3.5e-98;  
 Matches 213; Conservative 19; Mismatches 40; Indels 2; Gaps 2;

QY 1 MRGSEVLLMWLLVAVGGTEHAYRPGRRYACCPGKRTSGLPGAGAAICQPPCRNGSCVQP 60  
 Db 4 MWGSELLVAVFLVLAADGTEHAYRPSRRVCTVGISGSISETFVQRTYQPLTTCDGH 63  
 QY 60 RACSTYRTYRTAYRRSPGLAPRRYACCPGKRTSGLPGAGAAICQPPCRNGSCVQP 119

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 22, 2004, 12:42:40 ; Search time 45 Seconds  
(without alignments)  
1914.144 Million cell updates/sec

Title: US-09-978-191A-506  
Perfect score: 1505  
Sequence: 1 MRGSEVLLMWLLVAVGGT.....SEQISFLEBQLGSCSKKDS 273

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL 25:\*

1: sp\_archaea:\*

2: sp\_bacteria:\*

3: sp\_fungi:\*

4: sp\_human:\*

5: sp\_invertebrate:\*

6: sp\_mammal:\*

7: sp\_mhc:\*

8: sp\_organelle:\*

9: sp\_phase:\*

10: sp\_plant:\*

11: sp\_rhodent:\*

12: sp\_virus:\*

13: sp\_vertebrate:\*

14: sp\_unclassified:\*

15: sp\_xvirus:\*

16: sp\_bacteriaph:\*

17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1505	100.0	273	4	Q96EG0 Q96eg0 homo sapien
2	1504	99.9	273	4	Q9UH71 Q9uhf1 homo sapien
3	1180	78.4	278	11	Q9QXT5 Q9qxt5 mus musculus
4	1103.5	73.3	265	11	Q9DCP5 Q9dcp5 mus musculus
5	727.5	48.3	190	11	Q9JKW3 Q9jkw3 rattus norv
6	682.5	45.3	280	13	Q7ZX70 Q7zxt0 xenopus lae
7	435.5	28.9	293	11	Q35447 Q35447 mus musculus
8	423.5	28.1	293	4	Q99944 Q99944 homo sapien
9	420.5	27.9	293	4	Q8IV30 Q8iv30 homo sapien
10	277.5	18.4	512	5	Q9SRQ1 Q9srq1 drosophila
11	256	17.0	1574	11	Q88281 Q88281 rattus norv
12	252.5	16.8	558	4	Q9UFK6 Q9ufk6 homo sapien
13	251.5	16.7	553	4	Q9NZL7 Q9nzl7 homo sapien
14	251.5	16.7	553	4	Q8IUX8 Q8iux8 homo sapien
15	251.5	16.7	554	4	Q9NY67 Q9ny67 homo sapien
16	249.5	16.6	553	4	Q8NBV0 Q8nbv0 homo sapien

17	245.5	16.3	327	11	Q8BPM8 Q8bpm8 mus musculus
18	245.5	16.3	550	11	Q9JUZ5 Q9jjz5 mus musculus
19	225.5	15.0	544	13	Q8AVH7 Q8avh7 xenopus lae
20	223	14.8	561	11	Q91V88 Q91v88 mus musculus
21	218.5	14.5	578	11	Q91ZD3 Q91zd3 mus musculus
22	218	14.5	1664	5	Q9TVQ2 Q9tvq2 caenorhabdi
23	217.5	14.5	528	11	Q9CXD8 Q9cx8 mus musculus
24	198	13.2	747	11	Q8VHF4 Q8vhf4 mus musculus
25	198	13.2	1004	11	Q8CGH7 Q8cgh7 mus musculus
26	198	13.2	1034	11	Q8VHL7 Q8vhl7 mus musculus
27	197.5	13.1	592	11	Q91XL5 Q91xl5 mus musculus
28	196	13.0	1034	11	Q8VIX5 Q8vix5 mus musculus
29	194.5	12.9	590	11	Q8C088 Q8c088 mus musculus
30	193	12.8	609	11	Q923T5 Q923t5 mus musculus
31	191.5	12.7	814	13	Q7ZX15 Q7zxl5 xenopus lae
32	189	12.6	678	4	Q14393 Q14393 homo sapien
33	185.5	12.3	1600	11	Q8K4G0 Q8k4g0 mus musculus
34	185.5	12.3	1666	11	Q8K4G1 Q8k4g1 mus musculus
35	184.5	12.3	2906	11	Q9WUH9 Q9wuh9 rattus norv
36	184	12.2	673	11	Q61592 Q61592 mus musculus
37	184	12.2	674	11	Q99K57 Q99k57 mus musculus
38	183.5	12.2	648	13	Q7T3H4 Q7t3h4 brachydanio
39	183.5	12.2	784	11	Q8BM43 Q8bm43 mus musculus
40	183.5	12.2	816	11	Q8R417 Q8r417 rattus norv
41	183.5	12.2	819	11	Q8U0M5 Q8u0m5 mus musculus
42	183.5	12.2	858	11	Q8BM06 Q8bm06 mus musculus
43	182.5	12.1	815	4	Q96JS2 Q96js2 homo sapien
44	181.5	12.1	648	5	Q9VJU4 Q9vju4 drosophila
45	181.5	12.1	648	5	Q9NKK7 Q9nkd7 drosophila

#### ALIGNMENTS

##### RESULT 1

Q96EG0 PRELIMINARY; PRT; 273 AA.  
AC Q96EG0; AC Q96EG0;  
DT 01-DEC-2001 (TREMBLrel. 19, Created)  
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)  
DE Similar to NEUL protein.  
OS Homo sapiens (Human)  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OC NCBI\_TaxID:9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Ovary;  
RA Strausberg R.;  
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC012377; AAH12377.1; -;  
DR Genew; HGNC:20594; EGFL7.  
DR GO; GO:0005509; F:calcium ion binding; IEA.  
DR InterPro; IPR001152; Asx hydroxyl\_s.  
DR InterPro; IPR001881; EGF\_Ca.  
DR InterPro; IPR006209; EGF\_like.  
DR Pfam; PF00008; EGF; 2.  
DR SMART; SM00179; EGF\_CA; 1.  
DR PROSITE; PS00010; ASX HYDROXYL; 1.  
DR PROSITE; PS00022; EGF\_1; 1.  
DR PROSITE; PS01186; EGF\_2; 1.  
DR PROSITE; PS01187; EGF\_CA; 1.  
DR EGF-like domain.  
KW  
SQ SEQUENCE 273 AA; 29631 MW; 5AD0A4845D5B688 CRC64;

Query Match 100.0%; Score 1505; DB 4; Length 273;  
Best Local Similarity 100.0%; Pred. No. 1.8e-127;  
Matches 273; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MRGSEVLLMWLLVAVGGTEHAYRGRRCVAVRAHGDVPSFSFVORVVPFLTTCDGHR 60  
Db 1 MRGSEVLLMWLLVAVGGTEHAYRGRRCVAVRAHGDVPSFSFVORVVPFLTTCDGHR 60